The perception of speech sounds recorded within the uterus of a pregnant sheep

Scott K. Griffiths, W. S. Brown Jr., and Kenneth J. Gerhardt
Department of Communication Processes and Disorders, Institute for the Advanced Study of the Communication Processes, University of Florida, 335 Dauer Hall, Gainesville, Florida 32611

Robert M. Abrams
Department of Obstetrics and Gynecology, Department of Pediatrics, University of Florida, Gainesville, Florida 32611

Richard J. Morris
Department of Communication Disorders, Florida State University, 107 Regional Rehabilitation Center, Tallahassee, Florida 32306

(Received 16 June 1992; revised 24 February 1994; accepted 3 May 1994)

The intelligibility of speech stimuli recorded within the uterus of a pregnant sheep was determined perceptually using a group of untrained judges. The intrauterine sound environment of the ewe was intended to simulate that of a pregnant woman. Two separate lists, one of meaningful and one of nonmeaningful speech stimuli, were delivered through a loudspeaker to the side of the ewe and were simultaneously recorded with an air microphone located 15 cm from the flank and with a hydrophone previously sutured to the neck of the fetus. Perceptual test tapes generated from these recordings were played to 102 judges. The intelligibility of the phonemes recorded in the air was significantly greater than the intelligibility of phonemes recorded from the uterus. A male talker's voice was more intelligible than a female talker's voice when recorded from within the uterus, but not so when recorded in the air. An analysis of the feature information transmission from recordings inside and outside the uterus revealed that voicing information is better transmitted *in utero* than place or manner information.

PACS numbers: 43.71.Gv, 43.71.Es, 43.80.Jz

INTRODUCTION

The fetus is a dynamic organism which responds to and is influenced by the many acoustic signals generated both inside and outside its mother. The characterization of endogenous and externally generated sounds in the maternal abdomen and the influence that these sounds have on the fetus have been the subjects of recent research (Armitage *et al.*, 1980; Querleu *et al.*, 1981; Gelman *et al.*, 1982; Vince *et al.*, 1982; Abrams *et al.*, 1989; Gagnon, 1989; Granier-Deferre and Abrams, 1989; Gerhardt *et al.*, 1990). For example, the human fetus responds to sound during the last trimester of gestation (Birnholz and Benacerraf, 1983; Querleu *et al.*, 1988), and may possess the ability to discriminate prosodic and segmental differences in speech stimuli and develop memories for speech (Fifer and Moon, 1988). Moreover, it has been suggested that the fetus responds to changes in the intrauterine environment, learns by association, and retains prenatal experiences into postnatal life (Smotherman and Robinson, 1987).

It is clear from research on sheep and humans that the fetus is exposed to sounds from the external environment, especially low-frequency noise. Very little reduction in sound pressure is provided by the tissues and fluids of the maternal abdomen and uterus as external signals pass to the fetal head (Querleu *et al.*, 1981; Vince *et al.*, 1982; Gerhardt *et al.*, 1990).

The fetus receives sensory information generated from the maternal organism and from the mother's external environment. Querleu *et al.* (1981) reported that the auditory environment consists of: (1) a low-level basal noise, including maternal and placental vascular sounds with incidental maternal intestinal noises; (2) exterior noises which are partially absorbed; and (3) exterior voices and the mother's voice which are also partially absorbed and altered. Furthermore, when Querleu and his colleagues (1988) had their subjects reproduce what they could hear when listening to the voices of the pregnant woman and other male and female speakers recorded *in utero*, these subjects were able to recognize only 30% of 3120 tokens of French phonemes. The authors reported no significant differences between the perception of external male and female voices, and although internal, the mother's voice was not significantly more intelligible.

It is during the latter stages of fetal development, when the hearing mechanism is intact, that the human fetus may be most influenced by the sound environment inside and outside its mother (Spence and DeCasper, 1987). Maternal vocalizations are easily transmitted to the amniotic fluid and represent the most intense sound normally present in the uterus (Querleu *et al.*, 1981; Vince *et al.*, 1985; Gerhardt, 1989). Cues inherent in the speech of both the mother and external talkers may be perceived by the fetus, thus forming the basis for language acquisition (Kuhl, 1976; DeCasper and Fifer, 1980). Indeed, the newborn is not totally insensitive to different speech sounds. Eimas *et al.* (1971) were the first to show that human infants, as young as 1 month old, demonstrate categorical perception of synthetic stop consonant-

 0001-4966/94/96(4)/2055/9/$6.00 © 1994 Acoustical Society of America

Downloaded 16 Aug 2011 to 130.95.128.51. Redistribution subject to ASA license or copyright; see http://asadl.org/journals/doc/ASALIB-home/info/terms

vowel (CV) syllables differing in voice onset time. Electrophysiologic studies in neonates less than 24-h old have shown distinct patterns in the averaged bioelectric response which corresponded to categorical differences in consonant voicing (Molfese, 1977).

Little is known about the extent to which speech sounds, emitted by external talkers, may be distorted as the sound is transmitted into the fetal environment. The purpose of this study was to evaluate the intelligibility of externally generated speech utterances transmitted to and recorded at the fetal head *in utero*. A pregnant ewe was used to process the speech stimuli because the transmissibility of sound in this animal model (Armitage *et al.*, 1980; Vince *et al.*, 1982) is similar to that of humans during pregnancy (Querleu *et al.*, 1981).

I. METHODS

A. Surgery

A pregnant ewe carrying a fetus at 140 days gestational age was fasted for 24 h and anesthetized with halothane (2.5%) in oxygen (97.5%). After sterile preparation, the fetal head was delivered through a midline abdominal incision and a hysterotomy according to procedures detailed by Abrams *et al.* (1984) for long-term chronic preparation. A miniature hydrophone (Bruel and Kjaer model 8103, Marlborough, MA), used to detect intrauterine sound-pressure levels (dB SPL *re*: 20 μPa), was sutured to the neck just below the pinna on one side of the fetal head. The hydrophone had previously been calibrated with a pistonphone (Bruel and Kjaer model 4223) and gas sterilized.

Ampicillin, 500 mg, was introduced into both the amniotic fluid and the peritoneal cavity. The fetal head was repositioned in the intrauterine cavity and the uterus and abdominal wall were closed. The hydrophone cable was tunneled subcutaneously, exited at the maternal flank, and was stored in a pouch sutured to the skin. The midline abdominal skin wound was closed with clips and the ewe was taken off anesthesia, extubated, placed back in her pen, and observed until awake. Surgery and subsequent care of animals conformed to the guidelines approved by the University of Florida.

After a recovery period of several days, the ewe was led into a cart and positioned 1.8 m from a loudspeaker (Peavey HDH-2). The center of the loudspeaker was adjusted to the same height as the center of the ewe's flank. The sides of the cart were open from the ventral surface of the abdomen to above the back, thus minimizing sound interference. The acoustic signal produced by an electronic artificial larynx applied to the ewe's abdomen was used to locate the position of the hydrophone. This location was determined by monitoring for the maximum output of the hydrophone with an oscilloscope while moving the artificial larynx along the abdomen (Gerhardt *et al.*, 1988). The fetus was in a position such that the hydrophone was closer to one flank of the ewe, which was positioned toward the loudspeaker.

A calibrated air microphone (Bruel and Kjaer type 4165) was positioned 15 cm from the ewe's flank and in-line with the center of the loudspeaker. Speech stimuli from the loudspeaker were recorded with both the microphone and hydrophone.

B. Speech stimuli

The speech utterances were based on the Griffiths (1967) lists and nonsense items. Each stimulus item was presented in a carrier phrase, "Say the word ____." Three lists of 50 consonant-vowel-consonant (CVC) words from the Griffiths lists and 14 vowel-consonant-vowel (VCV) nonsense items were spoken by either a male or a female talker. The mean fundamental frequencies were 136 and 186 Hz for the male and female talkers, respectively. Six lists of 64 words each were recorded, three different lists by the male talker and three by the female talker.

Each list (64 items per list), spoken by both male and female talkers, was played through the loudspeaker via a cassette tape recorder at three different airborne levels measured at the maternal flank: 85, 75, and 65 dB SPL (dB *re*: 20 μPa). These levels were selected to represent loud and conversational speech levels. The outputs from the hydrophone and the air microphone were recorded on separate channels of an FM tape recorder (Bruel and Kjaer type 7006).

C. Perceptual testing

The recordings were used to construct a perceptual test of speech intelligibility. The test required groups of judges to listen to the utterances in the carrier phrase and mark on a paper what they heard. The judges' responses thus provided the basis for determining intelligibility scores (percent correct) associated with the CVC words and the VCV nonsense items.

For the 50 CVC words, each judge selected his or her response from a closed set of six one-syllable words which differed in either the initial or final consonant (see the Appendix). For example, one stimulus item was "Say the word bite" and the response list included "right, kite, bite, height, spite, night." To be correct, the judge would have to mark the word "bite."

For the 14 VCV nonsense items, the judges filled in a blank in a *V*____*V* frame with the vowel set to /ɑ/. For example, if a judge heard, "Say the word /ɑpɑ/," he or she would have to write a "p" in the blank to be correct.

Four different perceptual tapes were developed. Each tape contained recordings from a single talker (the male or the female) using different transducers (microphone or hydrophone) at three stimulus levels (see Table I). Each group of judges heard only one perceptual tape and, thus, only one talker speaking three lists. Each tape consisted of three different sets of 64 utterances from a single talker in descending order of sound-pressure level (one set of 64 utterances at 85 dB, one set at 75 dB, and one set at 65 dB). Two groups of judges heard tapes with a hydrophone recording first (tapes A and C), and two groups heard tapes with a microphone recording first (tapes B and D). For example, one group of 29 judges heard the male talker's voice recorded with the microphone at 85 dB, then with the hydrophone at 75 dB, and finally with the microphone at 65 dB (tape B). No judge ever heard the same stimulus more than once.

Downloaded 16 Aug 2011 to 130.95.128.51. Redistribution subject to ASA license or copyright; see http://asadl.org/journals/doc/ASALIB-home/info/terms.jsp

TABLE I. Design of perceptual tapes indicating sequence of recording site and level of presentation. Each judge heard only one of the four perceptual tapes.

Tape	Talker	Sequence			No. of judges
A	male	hydrophone 85 dB	microphone 75 dB	hydrophone 65 dB	20
B	male	microphone 85 dB	hydrophone 75 dB	microphone 65 dB	29
C	female	hydrophone 85 dB	microphone 75 dB	hydrophone 65 dB	26
D	female	microphone 85 dB	hydrophone 75 dB	microphone 65 dB	27

Perceptual testing was completed by 102 judges, with between 20 and 29 judges auditing each perceptual tape. All testing was conducted in a specially designed listening laboratory which accommodated up to 30 people at one time. All of the perceptual tapes were played over loudspeakers in the free field at an output level set to be heard comfortably by all judges.

In order to meet the variance assumptions for statistical analysis the percent intelligibility data for both the CVC and VCV stimuli were transformed using an arc-sine function prior to further analysis (Winer, 1971). The transformed data for each syllable position (initial, medial, final) were subjected to a three-way analysis of variance (ANOVA) (SPSS, Inc., 1988) for the effects of transducer site (at the maternal flank versus *in utero*), talker gender, and stimulus level (85, 75, and 65 dB SPL at the maternal flank). *Post hoc* analyses of the significant main effects and interactions were completed using Duncan's multiple range test.

Finally, error matrices were developed for the VCV nonsense stimuli. Sequential Information Analysis (SINFA; Wang, 1976) was applied to the error matrices to evaluate the amount of feature information received. SINFA allows for the partitioning of the contingent information transmitted and received for particular features of the stimuli (e.g., voicing, manner, place). From these results a relative measure of performance may be calculated (the ratio of the bits of information received to the bits sent, with the effect of other features held constant). Although SINFA was designed for the analysis of closed set data (subjects here performed in a free-response mode), its application here does yield contingent performance measures which can be compared across stimulus conditions. Only the data from the 85-dB stimuli (the first block from each perceptual tape) were analyzed using SINFA.

II. RESULTS

A. Intelligibility

The speech intelligibility scores (percent correct) derived from the judges' responses to the perceptual tapes for the CVC words and VCV utterances by gender and transducer site are displayed in Figs. 1 and 2, respectively. Summaries of the means and standard deviations for intelligibility by transducer site, talker gender, and stimulus level are presented in Tables II and III. The results of the ANOVA

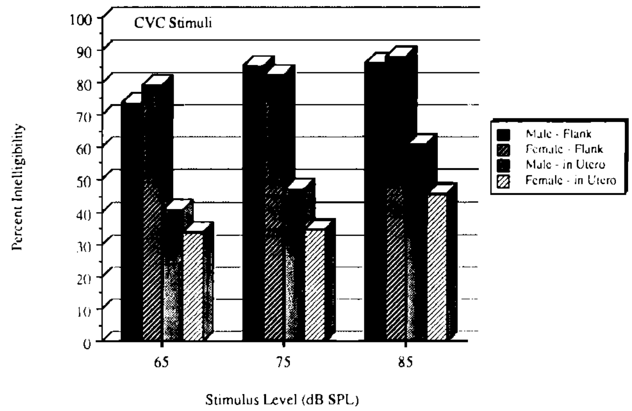

FIG. 1. Mean percent intelligibility of CVC words spoken by a male and a female talker recorded at the flank and within the uterus of a sheep at three airborne stimulus levels.

indicated significant main effects ($p<0.001$) for each of the three factors: transducer site ($F=3400.01$; $df=1$, 294), talker gender ($F=56.34$; $df=1$, 294), and level ($F=133.8$; $df=2$, 294). Although intelligibility scores were directly related to stimulus levels when recorded at the flank, the differences among the scores were slight. Moreover, there was no significant difference between the male and female talker, regardless of stimulus type when recordings were made at the flank.

Similarly, intelligibility scores recorded *in utero* improved with stimulus level for both stimulus types and both talkers. As can be seen in Figs. 1 and 2, these scores were significantly lower than those recorded at the flank. The data in these figures also reveal that the male talker was more intelligible than the female talker at all stimulus levels for both types of stimulus (by 21.6% on average). Thus, the male voice was more intelligible than the female voice when recorded *in utero* even though the intelligibility scores for both talkers were equivalent when recorded at the flank.

The mean intelligibility scores recorded *in utero*, averaged across the three levels and stimulus types (CVC and VCV) for the male and female voices, were 55.1% and 33.5%, respectively. At the flank, scores were approximately

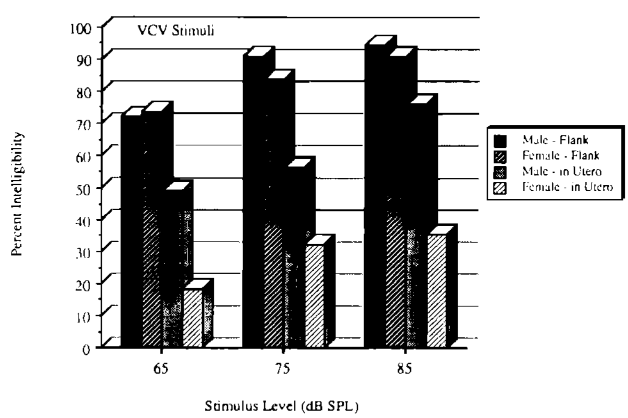

FIG. 2. Mean percent intelligibility of VCV nonsense stimuli spoken by a male and a female talker recorded at the flank and within the uterus of a sheep at three airborne stimulus levels.

Downloaded 16 Aug 2011 to 130.95.128.51. Redistribution subject to ASA license or copyright; see http://asadl.org/journals/doc/ASALIB-home/info/terms

TABLE II. CVC word intelligibility scores for each talker, transducer site, and stimulus level.

Male talker, CVC	85 dB	75 dB	65 dB
Microphone at flank			
Mean intelligibility	86.0%	85.3%	73.4%
No. correct/50	43	42.65	36.724
Standard deviation	2.104	2.084	2.776
No. of judges	29	20	29
Hydrophone *in utero*			
Mean intelligibility	60.7%	47.2%	40.9%
No. correct/50	30.35	23.586	20.45
Standard deviation	3.15	3.822	4.419
No. of judges	20	29	20
Female talker, CVC			
Microphone at flank			
Mean intelligibility	87.9%	82.5%	79.1%
No. correct/50	43.926	41.231	39.556
Standard deviation	1.73	2.355	2.979
No. judges	27	26	27
Hydrophone *in utero*			
Mean intelligibility	45.6%	34.7%	34.0%
No. correct/50	22.808	17.37	17
Standard deviation	2.315	2.256	3.464
No. of judges	26	27	26

83% for both the male and female talkers. Thus the reduction in intelligibility of recordings made from the flank site to recordings made *in utero* was 28.5% for the male speaker and 49.5% for the female speaker.

B. Effects of syllable position

A further analysis was completed on the target consonant sounds according to their position in the syllable: initial, medial, or final. Perception of all consonants in all positions was significantly ($p<0.01$) affected by stimulus level (F from 4.16 to 20.05; $df=2$, 612) and transducer site (F from 91.26 to 214.01; $df=1$, 612) with the 85-dB level exhibiting higher scores than the 75- and 65-dB levels. The overall intelligibility scores for the word final phonemes did not differ significantly from those for the medial phonemes. The scores for the word initial phonemes were significantly lower than the scores for phonemes in the other two positions ($F=7.69$; $df=2$, 610). This finding is consistent with the results of Griffiths (1967) in the moderately poor signal-to-noise ratio conditions (−4, −8 dB S/N). Griffiths suggested that these differences were spurious and should be ignored. While the intelligibility of word initial consonants was unaffected by talker gender, final consonants were correctly perceived significantly less often for the female than the male talker ($F=5.292$; $df=1$, 136). This effect was largely due to the poorer scores found for the *in utero* recordings of the female talker, as is demonstrated by the significant interaction between talker gender and transducer site ($F=10.756$, $df=1$,136).

TABLE III. VCV stimulus intelligibility scores for each talker, transducer site, and stimulus level.

Male talker, VCV	85 dB	75 dB	65 dB
Microphone at flank			
Mean intelligibility	94.3%	90.7%	71.9%
No. correct/14	13.207	12.7	10.069
Standard deviation	1.114	0.865	1.223
No. of judges	29	20	29
Hydrophone *in utero*			
Mean intelligibility	76.1%	56.4%	49.3%
No. correct/14	10.65	7.897	6.9
Standard deviation	1.496	1.372	2.222
No. of judges	20	29	20
Female talker, VCV			
Microphone at flank			
Mean intelligibility	90.7%	83.8%	73.8%
No. correct/14	12.704	11.371	10.333
Standard deviation	1.103	0.874	1.664
No. of judges	27	26	27
Hydrophone *in utero*			
Mean intelligibility	35.7%	32.5%	18.4%
No. correct/14	4.962	4.556	2.577
Standard deviation	1.685	1.502	0.987
No. of judges	26	27	26

C. VCV feature transmission

Consonant confusion matrices derived from the intervocalic (nonsense) stimuli for both talker genders and transducer sites for each of three levels are presented in Tables IV–VI. The open set nature of the nonsense item portion of the perceptual task allowed for more potential responses than there were stimulus items and for the possibility of the judges not responding. Because responses were written in English orthographics, it was impossible to know whether judges who wrote "th" intended the voiced or unvoiced linguadental fricative. For similar reasons judges were not able to uniquely indicate the voiced palatal fricative, /ž/, if they thought they perceived such a sound.

Given these limitations, several patterns still emerge from the sequential information analysis (SINFA). Table VII contains the percentage of contingent voicing, manner, and place information received (bits received/bits sent) for each talker and transducer site for the 85-dB stimuli.

A number of observations can be made from inspection of Table VII. Voicing information appeared to be well transmitted in recordings made at the flank regardless of talker gender. Voicing information received from *in utero* recordings was reduced for the female talker but not for the male talker. The place information lost by going from the flank to within the uterus was marked, especially for the female talker. Place information appeared to be better preserved for the male talker *in utero* (only dropping to 71% vs 34% for the female talker). Less information regarding manner of articulation was received *in utero* than in air with a larger reduction for the female speaker than the male.

III. DISCUSSION

Previous research on the transmission of sound pressure into the maternal abdomen and uterus has generated a consistent set of filter characteristics for external sounds at the ear of the fetus (Vince *et al.*, 1982; Gerhardt *et al.*, 1990).

wnloaded 16 Aug 2011 to 130.95.128.51. Redistribution subject to ASA license or copyright; see http://asadl.org/journals/doc/ASALIB-home/info/terms.

TABLE IV. Consonant confusion matrices for 85-dB stimuli.

MALE SPEAKER recorded AT FLANK

Stimulus: / Response:	p	b	t	d	k	g	s	z	ʃ	tʃ	f	v	m	n
p	25													
b		26										3		
t			29											
d				24										1
k	1				29									
g				4		29						2		
s							29							
z								29						
ʃ									29					
tʃ										29				
f											28	1		
v		2									1	19		
m		1										2	28	
n				1								1	1	28
θ												1		
h	3													
j														
l														
ŋ														
No Response:														

FEMALE SPEAKER recorded AT FLANK

Stimulus: / Response:	p	b	t	d	k	g	s	z	ʃ	tʃ	f	v	m	n
p	20													
b		27									1	1		
t			25											
d				10										
k	7		2		27									
g				17		27								
s							26				3			
z							1	27						
ʃ									27	3				
tʃ										24				
f											23			
v												26		
m													27	
n														27
θ														
h														
j														
l														
ŋ														
No Response:														

MALE SPEAKER recorded IN UTERO

Stimulus: / Response:	p	b	t	d	k	g	s	z	ʃ	tʃ	f	v	m	n
p	17										6			
b		19										7		
t			17		3									
d				14		1								
k	1		3		17					1				
g				6		19		17				3		
s							15		1	1				
z								2						
ʃ							1		18	6				
tʃ										11				
f							3				13			
v		1										9		
m													20	
n														20
θ											1			
h	2													
j								1	1	1				
l												1		
ŋ														
No Response:							1							

FEMALE SPEAKER recorded IN UTERO

Stimulus: / Response:	p	b	t	d	k	g	s	z	ʃ	tʃ	f	v	m	n
p	16		16		2		1			7	12	1		
b		19				5	1	3				9		
t	2		6		13		1	1	1	5	3			
d				6		1		2			1			
k	4		1		6			2	1	2				
g				12		10	1	5						
s							8		21	1				
z														
ʃ									1					
tʃ							1							
f		1				1	9		2	9	9			
v			1			1						8		
m				1		3						4	20	4
n				1		1		1					4	21
θ							1							
h	4	3	2		5	2	1			1	1	1		
j								2						
l				6		2		1				3	2	1
ŋ								1						
No Response:		3					2	8		1				

TABLE V. Consonant confusion matrices for 75-dB stimuli.

MALE SPEAKER recorded AT FLANK

Stimulus: / Response:	p	b	t	d	k	g	s	z	ʃ	tʃ	f	v	m	n
p	17													
b		19										12		
t			20											
d				15										1
k					20					1				
g				5		20								
s							20		1					
z								20						
ʃ									19					
tʃ										19				
f	1										19			
v		1										8		
m													20	
n														19
θ	1													
h	1										1			
j														
l														
ŋ														
No Response:														

FEMALE SPEAKER recorded AT FLANK

Stimulus: / Response:	p	b	t	d	k	g	s	z	ʃ	tʃ	f	v	m	n
p	12		4								7			
b		25	1									4		
t	8		2				2							
d				25										
k			2		26					1				
g				1		26		2						
s			1				24		1					
z								24						
ʃ									24					
tʃ										25				
f	6		12								19			
v		1										22		
m													26	
n														26
θ			2											
h			1											
j														
l			1											
ŋ														
No Response:									1					

MALE SPEAKER recorded IN UTERO

Stimulus: / Response:	p	b	t	d	k	g	s	z	ʃ	tʃ	f	v	m	n
p	24				1		3				22			
b		29		4		3						28		
t	2		19			1	1			6				
d				8		6		2	1					
k	1		10		25		8		1	1				
g				17		18		18	1					1
s							11		1	10				
z								7	1					
ʃ							5		21	8				
tʃ									1	4				
f	2						1				4			
v						1						1		
m													29	
n														28
θ														
h					3						3			
j								2	2					
l														
ŋ														
No Response:														

FEMALE SPEAKER recorded IN UTERO

Stimulus: / Response:	p	b	t	d	k	g	s	z	ʃ	tʃ	f	v	m	n
p	4				1	1	1			2	3			
b		21		9		3		12			3	17		
t	2		1		3		1		1	2				
d			1	2		7								
k	9		6		13		3		1					
g		3	5	14		13		5						
s	4		3		3		10	3	13	12	1			
z									1	2	1			
ʃ	2		1		1		2	1	5		2			
tʃ	3		1				1		5	1				
f	1		1		2		7		1	5	14			
v		3	1			1		4		1		7		
m			1	1	1			1				3	24	17
n			1								1		3	10
θ														
h	2		2		1		1				1			
j					1									
l						1								
ŋ														
No Response:			3	1	1	1	1	1		2	1			

wnloaded 16 Aug 2011 to 130.95.128.51. Redistribution subject to ASA license or copyright; see http://asadl.org/journals/doc/ASALIB-home/info/terms

TABLE VI. Consonant confusion matrices for 65-dB stimuli.

MALE SPEAKER recorded AT FLANK

Response: \ Stimulus:	p	b	t	d	k	g	s	z	ʃ	tʃ	f	v	m	n
p	24		1								9			
b		23									6	20	1	
t	3		8									1		
d				18				3						1
k			20		29									
g		1		10		29		17						5
s							28							
z							1	4						
ʃ									29					
tʃ										29				
f	1										14			
v		3		1				2				6		
m												2	28	
n		1												23
θ		1						2						
h	1													
j														
l														
ŋ														
No Response:								1						

FEMALE SPEAKER recorded AT FLANK

Response: \ Stimulus:	p	b	t	d	k	g	s	z	ʃ	tʃ	f	v	m	n
p	16	2			1						12			
b		21										3		
t			26		3				4	13				
d				6										
k	11		1		23									
g				21		25								
s							26		6		1			
z							1	26						
ʃ								1	11					
tʃ									5	14				
f		3									12			
v		1										19		
m													27	
n														27
θ														
h											1	5		
j						1								
l														
ŋ														
No Response:						1			1		1			

MALE SPEAKER recorded IN UTERO

Response: \ Stimulus:	p	b	t	d	k	g	s	z	ʃ	tʃ	f	v	m	n
p	14		2		2					1	7			
b		20		2		4		3			6	15	5	4
t	2		8		1		8			7				
d	1			10		4						1		
k			9		17		9			3				
g				8		12		11	1			1		2
s			1				3		1	1				
z								6						
ʃ									12	2				
tʃ									4	3				
f	1									1	7			
v													1	
m													14	2
n												1		11
θ														
h	2											2		1
j									2					
l														
ŋ														
No Response:										2				

FEMALE SPEAKER recorded IN UTERO

Response: \ Stimulus:	p	b	t	d	k	g	s	z	ʃ	tʃ	f	v	m	n
p	9	4	5		5	1	6		2	2	7		2	
b	2	4	6	2			1	6	1	1		7	1	
t	5	1	1		8		4		3	10	5			
d						2		1	1					
k	6	8	6	1	2	2	3	1	3	6	2			
g		2	1	3	1	3						2		
s		1			1		5		5	1				
z					1				1					
ʃ			1						1		1			
tʃ					1	1			3	1	1			
f		3	2		3			2	4		5		1	
v			1	1		1		5	1			1		1
m				2	1	1	3	7			1	11	16	6
n		1		10	2	5	1	1		1		4	4	19
θ														
h	4	2	3	6		7	3	1	1	3	3		2	
j														
l				1		3		1				1		
ŋ														
No Response:					1			1		1	1			

Sound pressures below 200 Hz pass through to the fetus unattenuated, and, in some cases, are enhanced. Above this frequency, sound is increasingly attenuated by up to 20 dB.

Sample spectrographic displays of one stimulus item recorded in four conditions (male and female talkers recorded in air and *in utero*) are shown in Fig. 3. The phrase spoken in each of the four spectrograms is "Say the word batch." The input gain of the spectrographic analysis equipment was increased for the *in utero* recordings; one can observe that the speech signal, while still detectable, is much closer in level to the ambient noise. The contrast between voiced and unvoiced portions of the phrase is apparent to some degree in all four spectrograms. The high-frequency noise associated with the release of the affricate, /ĉ/, is undetectable in the *in utero* spectrograms for both talkers, consistent with the low-pass filtering of the maternal tissues and fluids.

Miller and Nicely (1955) reported that low-pass filtering of speech signals resulted in a greater loss of manner and place information than of voicing information. The authors concluded that the higher frequency information in the speech signal is critical for accurate identification of manner and place of articulation. In the current study, SINFA results indicated that the judges accurately perceived the voicing information *in utero*. Manner and place information were not transmitted as well as voicing information in the intrauterine recordings. These results could be predicted from the findings of Miller and Nicely given that transmission into the uterus can be modeled as a low-pass filter (Gerhardt *et al.*, 1990; Peters *et al.*, 1993). The poorer *in utero* reception of place and manner information associated with higher frequency sound pressures is consistent with the greater high-frequency attenuation.

TABLE VII. Conditional percentage of voicing, manner, and place information received (of bits sent) for both talkers and transducer sites in the 85-dB condition.

Talker:	Male		Female	
Site:	in air	*in utero*	in air	*in utero*
Information				
Voicing	95.5	100	95.8	79.5
Place	92.8	70.6	95.3	34.3
Manner	87.6	66.4	81.1	39.3

Voicing information from the male talker, which is carried by low frequencies, was largely preserved *in utero*. In fact, the judges perceived the male talker's voice equally well regardless of transducer site. Voicing information from the female carried less well into the uterus. The fundamental frequency of the female talker was approximately one half-octave higher than that of the male talker. Thus it is understandable that voicing information from the male talker would carry better into the uterus than that from the female talker.

Transmitted place and manner of articulation information from both talkers was significantly lower for the *in utero* site than at the flank. In general, for the *in utero* site, the judges made more errors on consonants that are considered to have greater concentrations of higher frequency energy (Miller and Nicely, 1955). This trend may be seen in comparing the *in utero* error patterns for fricative sounds to those for the nasal sounds in Tables IV–VI. This finding also re-

Downloaded 16 Aug 2011 to 130.95.128.51. Redistribution subject to ASA license or copyright; see http://asadl.org/journals/doc/ASALIB-home/info/terms.jsp

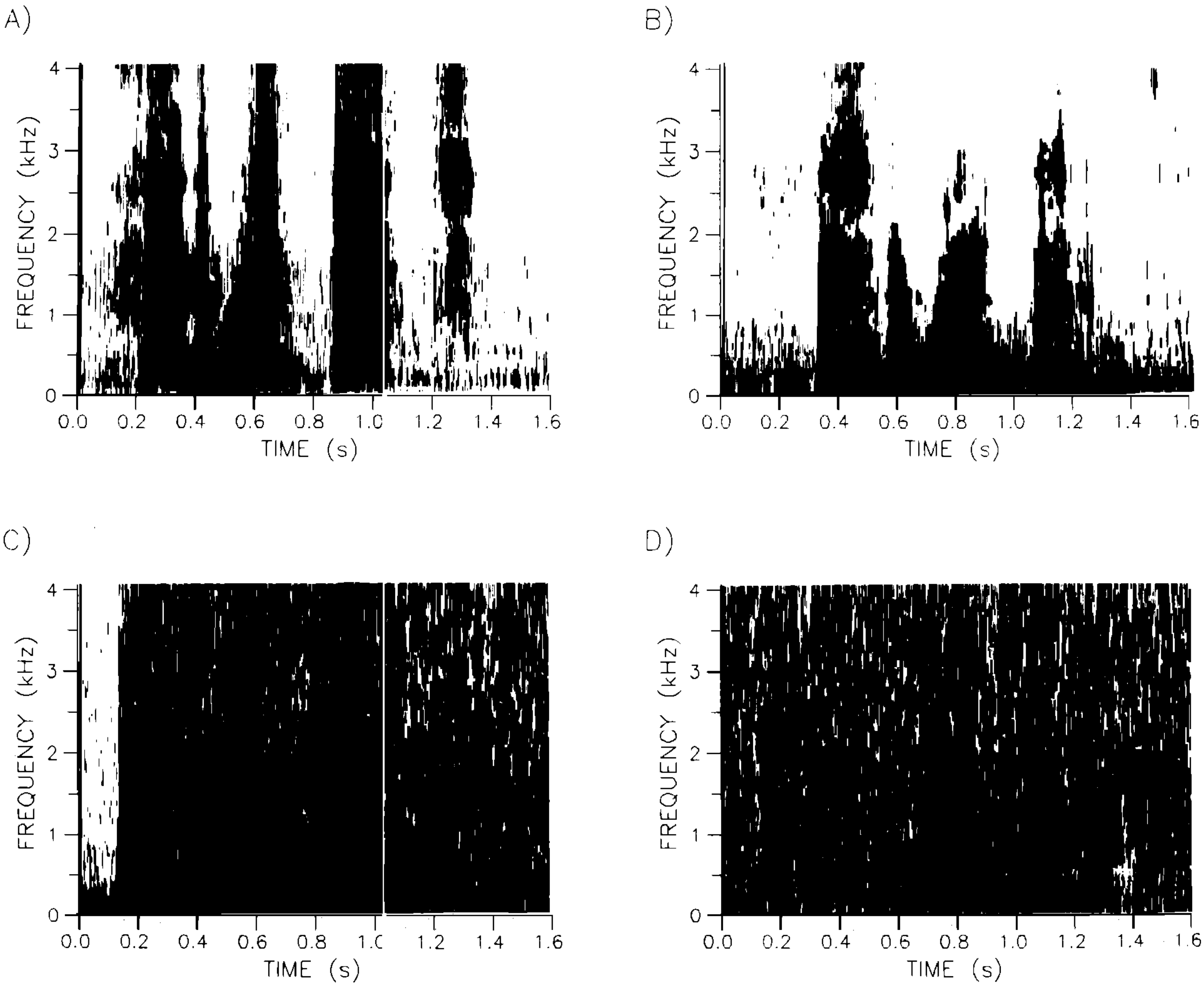

FIG. 3. Spectrographic recordings of "Say the word batch" spoken by: (a) the male talker recorded in air; (b) the female talker recorded in air; (c) the male talker recorded *in utero*; and (d) the female talker recorded *in utero*.

inforces the low-pass filtering of the maternal tissues and intrauterine fluids.

The design selected for this study did not include randomization of stimulus levels, although it did include randomization of the speaker and recording site conditions. All judges heard the 85-dB level first and, consequently, any learning effects were likely to be established at this level. Because the potential exists for judges to benefit from listening experience with the higher level stimuli, perceptual scores at the lower levels might be higher than if the stimulus levels were completely randomized. If stimulus levels had been randomized, some judges would have heard the 85-dB level first and potentially learned from that, while other judges would have heard a more difficult listening level first and learning may not have occurred until a later stimulus level.

The present findings represent better intelligibility for speech *in utero* than has been previously found (Querleu *et al.*, 1988). This difference occurs in spite of the fact that both sets of data were obtained using a combination of meaningful CVCs and nonmeaningful stimuli spoken by male and female talkers. A potential explanation for this discrepancy is a difference in the location of the transducer. The transducer in Querleu's study was positioned at the crown of the fetal head, potentially closer to vascular beds and therefore better able to pick up maternal heart sounds. D. J. Richards (personal communication) found no heart sounds when the transducer is passed beyond the crown of the human fetal head. The present study used a hydrophone sutured to the fetal neck. The absence of detectable heart sounds in the present recordings supports the report that this placement results in less vascular noise.

The implications of prenatal auditory discrimination and learning have attracted the attention of language scientists who have hypothesized that infants upon birth are equipped with either general sensory or specific linguistic mechanisms designed for the perception of speech (Jusczyk, 1993; Eimas *et al.*, 1971). Some theorists hold that human infants are born with a "speech module" (Liberman, 1982; Fodor, 1983), a mechanism designed specifically for processing the complex and intricate acoustic signals needed by humans to communicate with one another. An alternative theory of the neonate's initial state suggests that infants enter the world without specialized mechanisms dedicated to speech and

Downloaded 16 Aug 2011 to 130.95.128.51. Redistribution subject to ASA license or copyright; see http://asadl.org/journals/doc/ASALIB-home/info/terms

language, but rather respond to speech using general sensory and cognitive abilities (Kuhl, 1987).

While the results of this study reflect the perceptibility of the speech energies present in the amnion, they do not specify what speech energy might be present at the fetal inner ear. Measures of acoustic transmission to the fetal inner ear are quite limited at present (Gerhardt *et al.* 1992). Much work needs to be completed before conclusions can be drawn regarding what speech energies reach and are able to be perceived by the fetus.

IV. SUMMARY AND CONCLUSIONS

From the present data we conclude the relatively high intelligibility of externally generated speech recorded in the uterus of a sheep. The data further support better transmission of speech information into the uterus for a male speaker than a female speaker. Finally, the intelligibility data reveal better preservation of voicing information inside the uterus than of place or manner information, with a greater loss of each variety of information for the female speaker than the male speaker.

The present work also delineates the type of phonetic information that is available in the intrauterine environment. Consistent with the low-pass filtering by maternal tissues and fluids of externally generated sounds, voicing information is available *in utero*, while speech energy conveying place and manner information is attenuated. A potential effect of the present results is to call into question the assumption that the neonate is devoid of experience with speech. Male and female talker intelligibility scores developed from *in utero* recordings averaged approximately 55% and 34%, respectively, when recorded from within the uterus. While these data reflect speech energies present *in utero*, they do not necessarily represent the perceptibility of speech by the fetus. A definition of both the acoustic coupling of the fetal skull to maternal tissues and fluids, and the transmission pathway to the fetal inner ear of fluid-borne acoustic energy is required before such predictions can be reasonably made.

ACKNOWLEDGMENTS

The authors are especially grateful for the able assistance of Tom Sawallis, Jenifer Dutka, Jacquelyn W. Combs, and Aemil J. M. Peters. The work reported in this paper was supported by NIH Grant No. HD20084.

APPENDIX: SUBJECT RESPONSE SHEET

1	2	3	4	5
bass	loss	wick	duff	cup
batch	laws	with	duth	cub
badge	lodge	wit	dumb	cud
bat	log	wig	duv	come
bash	long	witch	dub	cuff
back	lob	will	dug	cut

6	7	8	9	10
dim	dung	fit	leash	toss
did	duv	fib	leave	talks
dill	dug	fig	liege	tall
dip	dud	fill	leach	tog
dig	done	fin	leed	tong
din	dub	fizz	leap	taj

11	12	13	14	15
lag	man	base	pan	peach
lash	mat	bays	pass	peas
lath	mad	bayed	pack	peal
lack	mack	beige	path	peat
lass	mass	bake	pad	peak
laugh	math	bathe	pat	peace

16	17	18	19	20
pitch	pus	has	weave	sash
pip	putt	hag	wean	sack
pig	puff	have	week	sad
pick	puck	half	weed	sap
pill	pup	hath	we're	sag
pit	pub	hash	weal	sat

21	22	23	24	25
sheath	sin	sud	tam	teer
sheave	sill	sup	tag	teeth
sheaf	sip	sub	tap	teethe
sheik	sick	sum	tang	teel
sheathe	sing	sun	tan	tease
sheen	sit	sung	tab	team

26	27	28	29	30
red	sold	wig	thick	tin
wed	hold	rig	chick	kin
dead	cold	gig	kick	fin
led	told	big	lick	shin
shed	gold	pig	sick	thin
fed	mold	dig	pick	pin

31	32	33	34	35
mark	tale	feel	till	peal
park	gale	eel	kill	zeal
dark	male	peel	hill	feel
bark	bale	keel	mill	reel
lark	pale	reel	will	veal
shark	rail	heel	bill	seal

36	37	38	39	40
same	then	fin	chin	zee
tame	ten	win	gin	thee
shame	fen	pin	tin	dee
game	hen	din	sin	knee
lame	den	sin	shin	see
came	pen	tin	thin	lee

41	42	43	44	45
tent	rip	shop	yore	fie
pent	lip	pop	for	thy
bent	chip	top	gore	vie
dent	tip	lop	wore	lie
rent	dip	cop	roar	thigh
went	hip	hop	lore	high

wnloaded 16 Aug 2011 to 130.95.128.51. Redistribution subject to ASA license or copyright; see http://asadl.org/journals/doc/ASALIB-home/info/terms.

46	47	48	49	50
dip	nest	rust	rat	may
zip	west	gust	mat	they
gyp	best	bust	bat	gay
ship	rest	lust	vat	bay
nip	jest	just	fat	nay
lip	vest	dust	that	way

(1) a____a	(2) a____a	(3) a____a	(4) a____a
(5) a____a	(6) a____a	(7) a____a	(8) a____a
(9) a____a	(10) a____a	(11) a____a	(12) a____a
(13) a____a	(14) a____a		

Abrams, R. M., Hutchison, A. A., and Gerhardt, K. J. (**1989**). "Effect of High Intensity Sound on Local Cerebral Glucose Utilization in Fetal Sheep," Dev. Brain Res. **48**, 1–10.

Abrams, R. M., Ito, M., Frisinger, J. E., Patlak, C. S., Pettigrew, K. D., and Kennedy, C. (**1984**). "Local Cerebral Glucose Utilization in Fetal and Neonatal Sheep," Am. J. Physiol. **246**, R608–R618.

Armitage, S. E., Baldwin, B. A., and Vince, M. A. (**1980**). "The Fetal Sound Environment of Sheep," Science **208**, 1173–1174.

Birnholz, J. C., and Benacerraf, B. R. (**1983**). "Development of Human Fetal Hearing," Science **222**, 515–518.

DeCasper, A. J., and Fifer, W. P. (**1980**). "Of Human Bonding: Newborns Prefer Their Mothers' Voices," Science **208**, 1174–1176.

Eimas, P. D., Siqueland, E., Jusczyk, P., and Vigorito, J. (**1971**). "Speech Perception in Infants," Science **171**, 303–306.

Fifer, W. P., and Moon, C. (**1988**). "Psychobiology of Newborn Auditory Preferences," Semin. Perinatol. **13**, 430–433.

Fodor, J. A. (**1983**). *The Modularity of the Mind* (MIT, Cambridge, MA).

Gagnon, R. (**1989**). "Stimulation of Human Fetuses With Sound and Vibration," Semin. Perinatol. **5**, 371–379.

Gelman, S. G., Wood, S., Spellacy, W. N., and Abrams, R. M. (**1982**). "Fetal Movements in Response to Sound Stimulation," Am. J. Obstet. Gynecol. **143**, 484–485.

Gerhardt, K. J. (**1989**). "Characteristics of the Fetal Sheep Sound Environment," Semin. Perinatol. **13**, 362–370.

Gerhardt, K. J., Abrams, R. M., Kovaz, B. M., Gomez, K. J., and Conlon, M. (**1988**). "Intrauterine Noise Levels in Pregnant Ewes Produced by Sound Applied to the Abdomen," Am. J. Obstet. Gynecol. **159**, 228–232.

Gerhardt, K. J., Abrams, R. M., and Oliver, C. C. (**1990**). "Sound Environment of the Fetal Sheep," Am. J. Obstet. Gynecol. **162**, 282–287.

Gerhardt, K. J., Otto, R., Abrams, R. M., Colle, J. J., Burchfield, D. J., and Peters, A. J. M. (**1992**). "Cochlear Microphonics Recorded from Fetal and Newborn Sheep," Am. J. Otolaryngol. **13**, 226–233.

Granier-Deferre, C., and Abrams, R. M. (**1989**). "Effects of Sound on Fetal Cerebral Glucose Utilization," Semin. Perinatol. **5**, 371–379.

Griffiths, J. D. (**1967**). "Rhyming Minimal Contrasts: A Simplified Diagnostic Articulation Test," J. Acoust. Soc. Am. **42**, 236–241.

Jusczyk, P. W. (**1993**). "Discovering Sound Patterns in the Native Language," in *The Proceedings of the Fifteenth Annual Conference of the Cognitive Science Society*, Boulder, CO, June, 1993 (unpublished).

Kuhl, P. K. (**1976**). "Speech Perception in Early Infancy: The Acquisition of Speech-sound Categories," in *Hearing and Davis: Essays honoring Hallowell Davis*, edited by S. K. Hirsh, D. H. Eldredge, I. J. Hirsh, and S. R. Silverman (Washington, U.P., St. Louis), pp. 265–280.

Kuhl, P. K. (**1987**). "Perception of Speech and Sound in Early Infancy," in *Handbook of Infant Perception, Volume 2: From Perception to Cognition*, edited by P. Salapatek and L. Cohen (Academic, New York), pp. 275–382.

Liberman, A. M. (**1982**). "On the Finding that Speech is Special," Am. Psychol. **37**, 148–167.

Miller, G. A., and Nicely, P. E. (**1955**). "An Analysis of Perceptual Confusions Among Some English Consonants," J. Acoust. Soc. Am. **27**, 338–352.

Molfese, D. L. (**1977**). "Infant Cerebral Asymmetry," in *Language Development and Neurological Theory*, edited by S. Segalowitz and F. Gruber (Academic, New York).

Peters, A. J. M., Gerhardt, K. J., Abrams, R. M., and Longmate, J. A. (**1993**). "Three-dimensional Intraabdominal Sound Pressures in Sheep Produced by Airborne Stimuli," Am. J. Obstet. Gynecol. **169**, 1304–1315.

Querleu, D., Renard, X., and Crepin, G. (**1981**). "Perception Auditive et Reactivite Foetale aux Stimulations Sonores," J. Gynecol. Obstet. Biol. Reprod. **10**, 307–314.

Querleu, D., Renard, X., Versyp, F., Paris-Delrue, L., and Vervoort, P. (**1988**). "La Transmission Intra-amniotique des Voix Humaines," Rev. Fr. Gynecol. Obstet. **83**, 43–50.

Smotherman, W. P., and Robinson, S. R. (**1987**). "Psychobiology of Fetal Experience in the Rat," in *Perinatal Development: A Psychobiological Perspective*, edited by N. A. Krasnegor, E. M. Blass, M. A. Hofer, and W. P. Smotherman (Academic, Orlando), pp. 39–60.

Spence, M. J., and DeCasper, A. J. (**1987**). "Prenatal Experience With Low Frequency Maternal-voice Sounds Influence Neonatal Perception of Maternal Voice Samples," Infant Behav. Dev. **10**, 133–142.

SPSS Inc. (**1988**) "SPSS/PC+V2.0," Author, Chicago, IL.

Vince, M. A., Armitage, S. E., Baldwin, B. A., *et al.* (**1982**). "The Sound Environment of the Foetal Sheep," Behavior **81**, 296–315.

Vince, M. A., Billing, A. E., Baldwin, B. A., Toner, J. N., and Weller, C. (**1985**). "Maternal Vocalizations and Other Sounds in the Fetal Lamb's Sound Environment," Early Hum. Dev. **11**, 179–190.

Wang, M. D. (**1976**). "SINFA: Multivariate Uncertainty Analysis for Confusion Matrices," Behav. Res. Methods Instrum. **8**, 471–472.

Winer, B. J. (**1971**). *Statistical Principles in Experimental Design*, 2nd ed. (McGraw-Hill, New York).

Downloaded 16 Aug 2011 to 130.95.128.51. Redistribution subject to ASA license or copyright; see http://asadl.org/journals/doc/ASALIB-home/info/terms