PHONETICS LABORATORY
LUND UNIVERSITY

WORKING PAPERS
4·1971

Previous numbers and contents:

Working Papers are distributed free of charge to those interested. Address:

Department of Phonetics
Lund University
Kävlingevägen 20
S-222 40 Lund
Sweden

WP 4·1971 printed September 1971

HAT DAS DEUTSCHE ZUGRUNDELIEGENDE STIMMHAFTE SPIRANTEN?

Untersuchung, initiiert durch W.U.Wurzels "Studien zur deutschen Lautstruktur" in Studia Grammatica VIII.

Robert Bannert

GRÜNDE DER FRAGESTELLUNG

Wurzels Studien zur deutschen Lautstruktur sind ein erster, begrüssenswerter Versuch, die phonologische und vor allem morphologische Seite des Deutschen im Sinne der generativen Grammatik nach Chomsky und Halle (1968) zu beschreiben. Die vorliegende Diskussion ist das Ergebnis einer spontanen Reaktion auf die Behandlung der stimmhaften Spiranten in Wurzels Liste des Systems der zugrundeliegenden Konsonanten im Deutschen (Wurzel, 248). Deshalb ist sie nicht als Gesamtkritik von Wurzels Studien zu betrachten. Sie beschäftigt sich nur mit den von ihm aufgestellten stimmhaften Spiranten /z/ und (/ɣ/). Alles andere der Studien bleibt deshalb - wenigstens einstweilen noch und wenn nicht ausdrücklich anders erwähnt - stehen.

Im folgenden versuche ich, eine andere und vielleicht plausiblere Darstellung der Spiranten im System der zugrundeliegenden Konsonanten des Deutschen zu geben als Wurzel. Wenn man die Aufstellung des Konsonantensystems auf S.248 betrachtet, fallen einem unwillkürlich zwei Dinge auf: zum ersten fehlt die stimmhafte labiale Spirans /v/, während /z/ einen Platz einnimmt und zum zweiten steht /ɣ/ in Klammern. Warum diese Asymmetrie bei den stimmhaften Spiranten?

p	t	k
b	d	g
f	s	x
-	z	(ɣ)

An verschiedenen Stellen im Text finden sich Hinweise auf diese Tatsachen:

1. "Das durch die Eliminierung eines zugrundeliegenden [v] entstehende "Loch" im System könnte zwar "unschön" erscheinen, belastet aber, jedenfalls im konkreten Fall, in keiner Weise die Generalität der Grammatik" (S.246).
2. Bei [z] sei keine "explizite Behandlung ihrer Problematik notwendig", während die palatale und die labiale Spirans eine von den anderen Spiranten unterschiedliche Stellung innerhalb des Systems der deutschen Konsonanten einnehmen (S.244).
3. "Das zugrundeliegende Segment /ɣ/ unterscheidet sich von allen anderen bisher behandelten durch sein sehr geringes Vorkommen. Es nimmt im deutschen Konsonantensystem eine besondere Stellung auch deshalb ein, weil es nur postvokalisch, nicht aber prävokalisch, d.h. also im Anlaut stehen kann" (S.240).

Selbst wer kein Anhänger des Symmetrieprinzips innerhalb der Phonologie ist, muss zumindest leise Zweifel an Wurzels Behandlung der Spiranten hegen; denn die Argumente für die gelieferte Darstellung sind keineswegs einleuchtend. Eine eingehende Überprüfung der Darstellung sollte Antwort geben auf folgende Fragen:

1. Warum fehlt die labiale Spirans /v/, während /z/ seinen Platz behauptet?
2. Warum steht /ɣ/ in Klammern?
3. Warum können nicht alle stimmhaften Spiranten als zugrundeliegende Segmente fehlen, wenn /v/ es kann?

Bereits Ross (S.51) vermutet, dass zugrundeliegendes /z/ überflüssig sei. Er schlägt /ɣ/ nur als Versuch vor und äussert gleichzeitig starke Zweifel an der Richtigkeit dieser Entscheidung.

DIE LABIALE SPIRANS [v]

Die Behandlung und Herleitung von [v] aus /w/ (Wurzel, 244f) ist einleuchtend, wenn auch z.T. die historischen Gesichtspunkte und der Vergleich mit anderen germanischen oderindoeuropäischen Sprachen zur Illustration und Argumentation gute Dienste hätten leisten können. Somit ist ein "Loch" im System gerechtfertigt, nämlich das fehlende /v/. Ob dieses "Loch" nun "unschön" ist oder nicht, d.h. ob die Reihe der stimmhaften Spiranten nur ein "Loch" aufweist oder vielleicht nur aus "Löchern" besteht, wird im folggnden untersucht.

DIE DENTALE SPIRANS [z]

Wie verhält es sich mit dem /z/? Warum findet Wurzel es nicht einmal der Mühe wert, dieses Problem zur Diskussion aufzunehmen? Als einzige Tatsache wird die "Auslautverhärtung" des /z/ an verschiedenen Stellen (z.B. S.245, 259f) angeführt. Ist diese Behandlung des /z/, parallel zur Neutralisation von /b,d,g/ am Wortende, richtig?

Auf den ersten Blick sieht es so aus, als ob das [z] tatsächlich "auslautverhärtet" auftritt:

[z]	[s]
Häuser	Haus
Nase	Nas'
lesen	lies

Genau so wie z.B.	b – p:	Leiber	– Leib
	d – t:	Bäder	– Bad
	g – k:	Tage	– Tag

Dieser Auffassung aber widerspricht die Beobachtung, dass das s der Schrift am Wortanfang vor Vokal aller Wörter ausnahmslos als [z] ausgesprochen wird: z.B. sagen, Seil, Sehn-sucht. In dieser Hinsicht verhält sich das

[z] nicht wie die stimmhaften Verschlusslaute, da die Opposition stimmhafter~stimmloser Verschlusslaut am Wortanfang häufig auftritt:

backen - packen
Diele - Thiele
Gunst - Kunst

Das stimmhafte s tritt immer am Wortanfang vor Vokal auf, gerade dort, wo eben kein stimmloses s möglich ist. [z] ist also in dieser Stellung stets vorauszusagen. Deshalb dürfte der Begriff "Auslautverhärtung" nicht für das [z] gelten.

Ein kurzer Seitenblick auf die historische Entwicklung (Schirmunski, 357f) zeigt unmittelbar, dass dieses germanische stimmlose s in der neuhochdeutschen Literatursprache am Wortanfang vor Vokal stimmhaft wurde, während es im Wortauslaut seine Stimmlosigkeit bewahrte. Die Entwicklung sieht also so aus:

(1) s wird z im Anlaut vor Vokal,
(2) s bleibt s im Wortauslaut.

Im Neuhochdeutschen wurde das germanische s gleichzeitig noch von einer weiteren Änderung betroffen. Es wird auch im Inlaut vor Vokal stimmhaft realisiert, wobei der vorhergehende Laut ebenfalls stimmhaft sein muss: z.B. lesen, Wiese, Hirse, Bremse, Sense. Ist der vorhergehende Laut demgegenüber stimmlos, bleibt s auch stimmlos: z.B. knipsen, Schnäpse, Achse, Hexe. In den Verbindungen -sp- und -st- bleibt s erhalten: z.B. Wespe, Kasten, Fenster, Elster. Damit sind (1) und (2) zu ergänzen:

(3) s wird z (a) im Anlaut vor Vokal,
(b) im Inlaut vor Vokal, wenn der vorhergehende Laut stimmhaft ist.

(4) s bleibt s (a) im Wortauslaut,
(b) vor oder nach stimmlosem Laut.

Sehr aufschlussreich in diesem Zusammenhang ist die Tatsache, dass diese

Entwicklung s ⟶ z in bestimmten Umgebungen (3) nicht für alle deutschen Dialekte gilt. Während z.B. im Niederländischen diese lautliche Entwicklung ihren Niederschlag sogar in der Schreibweise gefunden hat (stimmhaftes s wird durch das Graphem z wiedergegeben), kennen die meisten oberdeutschen Dialekte ebensowenig stimmhaftes s wie stimmhafte Verschlusslaute (Schirmunski, 358f).

Es ist leicht zu sehen, dass das s überall da, wo es zwischen zwei stimmhaften Lauten zu stehen kommt, ebenfalls stimmhaft wird. Offenbar handelt es sich um eine gleichzeitig regressive und progressive Assimilation, wobei der vorhergehende stimmhafte Laut Voraussetzung für das Eintreffen der Anpassung des Merkmals "stimmhaft" des folgenden Lautes ist.

Bevor nun diese Erscheinung zu formulieren ist, muss geprüft werden, ob sie eine Morphemstrukturbedingung des Deutschen darstellt (vgl. Stanley, 1967), also ins Lexikon gehört, oder als phonologische, wahrscheinlich späte Regel zu bestimmen ist.

Es folgt dazu eine Zusammenstellung von Wörtern, die das zur Diskussion stehende s enthalten. Gegebenenfalls sind diese Wörter in Morpheme analysiert: Morphemgrenze +, einfache Wortgrenze # (Wurzel, 249):

(5) #sand   treib# #sand
     #sing   sitt# #sam

(6) (a)   lēs + (e)n   läus + e
     (b)   nās#   weis#

(7)   ēsl   amsl   pinsl

In (5) wird s nach # vor Vokal als [z] realisiert. Diese Feststellung gilt ausnahmslos für alle als deutsch, d.h. nicht einer bestimmten Fremdsprache angehörend, empfundenen Wörter, also auch Lehn- und Fremdwörter wie sozial, Sofa, Samt usw. Im Wortauslaut, also vor #, steht [s] (6b). In diesen Fällen wird [s] zu [z], sobald ein Derivativ folgt: z.B. nās + e, weis+e. In (6a) wird *s* als [z] realisiert, also vor + mit folgendem stimmhaften Laut. Die Beispiele unter (7) zeigen das als [z] realisierte *s* zwischen stimmhaften Segmenten innerhalb des Morphems.

Bis hierher sind die Zusammenhänge leicht zu überblicken. Eine Komplikation stellen die Suffixe -sal/-sel dar. Beide haben denselben Ursprung. Während -sal immer als [zāl] mit Nebenakzent erscheint, wird -sel entweder als

(8) [zəl] bzw. [zl̩]: Füll-sel, Streu-sel oder als

(9) [səl] bzw. [sl̩]: Rät-sel, Kap-sel. Überbleib-sel

stets unbetont realisiert. Das *s* in -sel erscheint demnach stets als [z], wenn das vorhergehende Morphem auf einen stimmhaften Laut endet (8), sonst als [s] (9).

Da -sel immer [-betont], kann nach dem Operieren der Akzentregeln (Wurzel, 272) keine ## davor stehen. Und da -sal heute kein selbständiges Wort ist, sondern nur in zweigliedrigen Zusammensetzungen vom Typ Schicksal vorkommt, trägt es nie den Hauptakzent. Es ist nach Kiparsky (1966, 72f) ein betonbares Suffix. Demgemäss erhält es z.B. in der Ableitung Schick-sal Nebenakzent, und die Ableitung das Akzentmuster Sch$\overset{1}{\text{i}}$ck-s$\overset{2}{\bar{\text{a}}}$l. Das Suffix -sel aber tritt in Ableitungen ausschliesslich tonlos auf, z.B. H$\overset{1}{\text{a}}$ck-s$\overset{4}{\text{e}}$l. Folglich muss -sel im heutigen Deutsch ein einfaches Wortbildungssuffix wie z.B. -ig oder -er darstellen. Es ist ein unbetonbares Suffix, obwohl -sel und -sal *historisch* gesehen identisch sind. Hier dürfte ein Fall von Umstrukturierung vorliegen (vgl. weiter unten S. 16). Die heutige, normale Realisierung des -sel ohne Vokal als [sl̩] bzw. [zl̩] bekräftigt diese Annahme. Die Parallele zu Formen wie Beutel, Onkel, Rasen, Wagen, die normalerweise mit silbischem [l̩] bzw. [n̩] realisiert werden, ist offenbar. Somit gelten für -sel die Regeln wie für z.B. wins+eln, Pinsel (vgl. Gerinn+sel) bzw. knipsen (vgl. Stöp+sel). Wurzels Analyse von *Geschreibsel* und *Überbleibsel*

(Wurzel, 260) mit der einfachen Wortgrenze vor -sel scheint deshalb nicht den Tatsachen zu entsprechen. Dies wiederum führt zu einer Vereinfachung seiner Regel der "Auslautverhärtung" (S.260). Die Stimmlosigkeit trifft nur stimmhafte Segmente im Auslaut, gilt aber nicht für das morpheminitiale s in -sel (Teil d der Regel K 13).

Ähnlich liegen auch die Verhältnisse bei dem Pronomen sie (Wurzel, 265). Solange es betont ist, steht die Wortgrenze vor s, das als [z] realisiert wird. Vor dem unbetonten enklitischen Pronomen ist diese Grenze weggefallen (Bierwisch, 1966), das s wird nach einem stimmlosen Segment als [s] realisiert. In diesem Fall liegen also dieselben Bedingungen vor, wie innerhalb des Morphems oder über die Morphemgrenze hinweg. Deshalb scheint Wurzel auch hier zu irren (S.266, 2.Absatz).

Diese rhythmische Verschmelzung in der Rede - in den Aussprachewörterbüchern unter dem Begriff "Umgangssprache" zu finden - trifft aber bei weitem nicht nur die Pronomen. Immer wenn die Sprache wirklich gesprochen wird - im Gegensatz zur reinen (genormten) Hochlautung - werden Grenzen in verschiedenem Masse gelöscht. Was das s betrifft, gerät es dann immer in dieselbe Position wie oben innerhalb des Morphems, da die Wortgrenze entfällt:

(10) in zusammengesetzten Wörtern.

Gras-samen	Gesichts-seife	Trüb-sal
Kreis-säge	Salz-säure	ab-sehen

(11) in Phrasen.

Jedem das Seine	Er hat sagen wollen
Gottes Segen	Wer mag solches
Was sehen sie	

Nach stimmlosem Segment operiert die Regel Stimmhaftigkeit des s nicht (siehe S.12). Die Regel Geminatenvereinfachung (siehe S.11) eliminiert das zweite /s/ des Doppelsegmentes /ss/ in den entsprechenden Beispielen.

Es erhebt sich nun die Frage der zugrundeliegenden Repräsentation der Morpheme mit s im Lexikon. Sind z.B.

Sand	als	/zand/	oder	/sand/,
-sal	"	/zal/	"	/sal/,
Maus	"	/mauz/	"	/maus/,
Esel	"	/ēzl/	"	/ēsl/

anzusetzen?

Zur Beantwortung dieser Frage ist eine Zusammenstellung von Beobachtungen notwendig:

s wird immer [z]:

1. nach der einfachen Wortgrenze # vor betontem Vokal sowie zwischen stimmhaften Segmenten innerhalb des Morphems,
2. nach stimmhaftem Segment vor Morphemgrenze mit anschliessendem stimmhaften Segment.

In allen anderen Fällen bleibt [s].

Allein auf grund dieser Regelmässigkeit scheint alles für die Repräsentation des realisierten [s] und [z] als /s/ zu sprechen, da die richtigen phonetischen Formen, also die Ausgabe der phonologischen Komponente der Grammatik, einfach mit Hilfe phonologischer Regeln erzeugt werden können. Eine besondere Morphemstrukturbediengung (MSB) würde den einheitlichen Prozess der Stimmhaftigkeit des s aufsplittern. Das entscheidende Argument gegen eine solche MSB ist die Tatsache, dass das stimmhafte s nur in einigen Dialekten und über diese auch in der hochdeutschen Schriftsprache realisiert wird, während andere Dialekte und ältere Sprachstufen kein stimmhaftes s kennen.

Bevor die Frage der Lexikoneintragung des s endgültig beantwortet werden kann, müssen die Fälle mit ß (scharfes s) und ss der Schrift, die immer stimmlos realisiert werden, untersucht und mit dem einfachen s koordiniert werden. Die folgende Tabelle enthält Wörter mit [s] und zum Vergleich solche mit [z]:

[s]		[z]
kurz.Vok./Diphth.	langer Vokal	lang.Vok./Diphth.
hasse		Hase
reißen		reisen
Risse, rissig		Riese, riesig
Masse	Maße	Masern
muß	Muße	Muse
fasse		Phase
wissen		Wiesen
weißen		weisen
Fussel	fußeln	Fusel
	Bußen	Busen

Wenn dieses stimmlose s ebenfalls durch ein einziges s-Segment im Lexikon repräsentiert würde, erscheinen als Ausgabe der phonologischen Komponente nicht richtige Formen wie *[štrāzə], *[fūzə] usw.

Die Annahme eines einfachen s-Segmentes in diesen Fällen liefert falsche Ergebnisse. Dieses s bleibt stimmlos zwischen stimmhaften Segmenten. Warum tut es das? Im Unterschied zu dem s, das stimmhaft wird, erscheint das stets stimmlose s als ss oder ß in der Schrift. Das ß ist nur eine orthographische Variante des ss, um in erster Linie bei intervokalischem stimmlosen s die Länge des vorhergehenden Vokals anzuzeigen, s.B. Straße – Gasse.

In fast allen Fällen handelt es sich um ss, das aus germanisch t nach Vokal entstand: Es ist ein Teilergebnis der zweiten hochdeutschen Lautverschiebung (Schirmunski, 272). Vgl. dazu folgende Zusammenstellung:

nhd.	(alt)ndd.	eng.	schwedisch
Wasser	watar	water	vatten
essen	etan	eat	äta
hassen		hate	hata
Fuß	fot	foot	fot
Füße		feet	fötter

Nun hat es natürlich keinen Sinn, ein zweites, vom erhaltengebliebenen germanischen s unterschiedliches s anzusetzen. Für den Sprecher des heutigen Deutsch ist das s in z.B. Haus und Hass ein und derselbe Laut (vgl. dazu weiter unten S.14 und Wurzel, 207).

Beim Betrachten der Liste auf S.9 fällt auf, dass das stimmlose s (ss oder ß der Schrift) nach kurzem oder langem Vokal steht oder nach Diphthong. Das einfache s steht nach langem Vokal oder Diphthong. Gibt der Vokal, gespannt oder ungespannt, Aufschlüsse über die Behandlung des s ? Oder umgekehrt, hängt das Merkmal [±gespannt] mit dem Merkmal [±stimmhaft] 1
des s zusammen? Folgende Tatsachen lassen sich feststellen:

1. Nach Diphthong und langem Vokal erscheinen [z] und [s].
2. Stimmloses s bleibt erhalten in den verschiedenen Formen der (zumeist) starken Verben sowie deren Nominalisierungen, während gespannter und ungespannter Vokal wechseln. Dieser Wechsel unterliegt einer phonologischen Regel (P 4 bei Wurzel, 75), z.B. flīssen-floss-geflossen, Fluss-Flüsse, messen-māss-gemessen, Māss-Māsse.
3. Dieses stimmlose s tritt nie wortinitial vor Vokal auf (bedingt durch seine Entstehung).

Daraus wird deutlich, dass die Behandlung des s in Abhängigkeit des vorhergehenden gespannten oder ungespannten Vokals im Deutschen nicht möglich ist. Dies offensichtlich im Gegensatz zum Englischen (Chomsky und Halle 1968, 46,149), wo das stimmhafte s am Wortanfang vor Vokal in einheimischen Wörtern nicht auftritt.

Zur Beantwortung der Frage der Lexikonrepräsentation des s muss noch eine (eigene) Beobachtung herangezogen werden. Wenn Deutsche eine Fremdsprache lernen, in der es nur die stimmlose dentale Spirans /s/ gibt, z.B. schwedisch, lässt sich normalerweise feststellen, dass – zumindest am Anfang – das einfache s der Schrift (a) am Wortanfang vor Vokal (sol) und (b) zwischen stimmhaften Lauten (visa, hals+en) als [z] realisiert wird, während alle ss richtig als [s] ausgesprochen werden (kasse), was auch für alle einfachen s am Wortende gilt (mås).

Das bedeutet, dass der Anfänger in der fremden Sprache die Regeln des Deutschen für die Behandlung des s konsequent auf die Fremdsprache anwendet. Da der Deutsche anfangs auch die Regeln der Stimmlosigkeit bei b,d,g auf die fremde Sprache anwendet, müssen die Regeln für die Stimmhaftigkeit des s ebenfalls aktiver Bestandteil der Grammatik des Deutschen sein.

Umgekehrt ist es z.B. für Schweden oft zumindest am Anfang schwierig, das einfache s im Deutschen korrekt zu realisieren. Nach Kitzings Untersuchung der Aussprache einiger schwedischer Schüler (1967) weist das [z] die bei weitem geringste Lösungsfrequenz unter den Spiranten und Affrikaten auf (10% initial in sonntags, 19% medial in Rose). Dagegen wird das [s] in Messer bzw. weiss zu 99 bzw. 100% richtig ausgesprochen.

Wenn das Schriftbild zusammen mit einer einfachen Regel die richtige Realisierung des s gewährleistet, scheint der Versuch gerechtfertigt zu sein, das [s] immer dann, wenn es zwischen stimmhaften Lauten zu stehen kommt (es kann ja nicht nach ## vor Vokal auftreten!), als /ss/ zu repräsentieren (vgl. Ross, 51). Das Doppelsegment wird später durch eine Eliminierungsregel vereinfacht, die bereits für zwei identische Segmente über Wort- und Morphemgrenze hinweg ihren Platz in der phonologischen Komponente hat (Wurzel, 221 ff). Ebenso wie verdoppelte Segmente über diese Grenzen hinweg vereinfacht werden, z.B. in Formen wie den Buben /būb+n+n/, er rät /rāt+t/, die durch Flexions- und Derivationsregeln auf bestimmten Stufen der Ableitung entstehen, werden im Fall des /ss/ auch zweifach repräsentierte Segmente zwischen Morphemgrenzen nur als eines realisiert. Damit ergibt sich ein weiterer empirischer Beweis für die Regel der Geminatenvereinfachung (Wurzel, 223). Indem sie auch innerhalb des Morphems operiert, gewinnt sie an Generalität.

Diese Lösung, die keine extra Merkmale kostet, liefert alle richtigen Formen. Gleichzeitig wird das s, das einfach repräsentiert wird, stets richtig auf grund phonologischer Regeln entsprechend seiner Umgebungen entweder als [s] oder [z] generiert.

Die Stimmhaftigkeit des einfachen s ist damit eine generelle Erscheinung, wonach es immer stimmhaft zwischen einfacher Wortgrenze bzw. stimmhaftem Segment und stimmhaftem Segment auftritt, ganz gleich, ob vor- oder nachher eine Morphemgrenze steht. Die Regel: Stimmhaftigkeit des s lässt sich zusammengefasst so formulieren:

$$\begin{bmatrix} +obstr \\ +dnd \\ +kor \end{bmatrix} \longrightarrow [+sth] \Big/ \left\{ \begin{matrix} \# \\ [+sth] \end{matrix} \right\} \underline{\qquad\qquad} [+sth]$$

An welcher Stelle in der phonologischen Komponente operieren die Regeln für s ? Die Lexikoneintragungen (nach Wurzel, ausser der Behandlung des s) sind z.B.

/fels+n/ /ūbr + blejb+sl/

/būss/ /xals/ /mows/

Die Beispiele zeigen, dass sich die Stimmhaftigkeit des s nur in den vollständigen Formen auswirkt. Durch die e-Epenthese-Regel (Wurzel, 171) entstehen die vollen Formen der Substantive (Genus, Kasus, Plural), wie auch die der Adjektive und Verben. Aber erst wenn auch die "Stimmlosigkeit" (Auslautverhärtung bei b,d,g) gewirkt hat, so dass z.B. ūberbleip+sel vorliegt, kann die Regel "Stimmhaftigkeit des s" arbeiten. Sie liefert dann alle korrekten Formen. Die Einordnung dieser Regel zeigt, dass sie eine späte phonologische Regel ist. Die Regelfolge (Wurzel, 288) dürfte sich wie folgt ändern:

⋮

(A 1) Ablaut bei starken Verben

⋮

(57) Analyse in phonologische Wörter

⋮

(SV 1) e-Epenthese

⋮

(SV 5) (e-Eliminierungen)

⋮

(K 13) Stimmlosigkeit b,d,g (Auslautverhärtung), ohne Teil (d)

(K 10) Wechsel vor t (wenn überhaupt in der Grammatik vorhanden)

Stimmhaftkeit des s

(K 5) Geminatenvereinfachung

⋮

DIE VELARE SPIRANS [ɣ].

Auf grund der Ergebnisse der vorhergehenden Diskussion bleiben folgende zugrundeliegende Spiranten übrig:

f	s	x
–	–	(ɣ)

Das (/ɣ/) steht allein in der stimmhaften Reihe. Wurzel braucht es einzig und allein, um 5 Wortpaare unter eine Regel (K 10) zwingen zu können, wobei er einfach den Vorschlag von Ross (S.51), gegen den dieser selbst Einwände erhebt, übernimmt:

(49)	fliehen	– Flucht, flüchten
	geschehen	– Geschichte
	sehen	– Sicht, Gesicht
	(ver)zeihen	– Verzicht
	ziehen	– Zucht

Für die Behandlung der Gruppen

(50)	hohe, höher, Höhe	– hoch, höchst
	nahe, näher, Nähe	– nach, nächst

(Hier ist das Paar: (ver)schmähen - Schmach, (Ross, 51), hinzuzufügen).

und (51) bringen, brachte, gebracht
denken, dachte, gedacht
mögen, mochte, gemocht

diskutiert Wurzel zwei Möglichkeiten, wobei für die zweite ebenfalls /ɣ/ gebraucht wird, ohne dass er aber eine endgültige Entscheidung trifft (Wurzel, 240f).

Im folgenden versuche ich zu zeigen, dass das Deutsche kein zugrundeliegendes /ɣ/ braucht. Aus der Diskussion dürfte hervorgehen, dass das Postulieren des /ɣ/ nur ein technischer Trick ist, einige, scheinbar ähnliche Zusammenhänge in ein Regelschema zu pressen.

Zuallererst können einige von Wurzels Behauptungen und Annahmen nicht unwidersprochen bleiben. Erstens sind die Prozesse

(39) ge<u>b</u>en ⟶ Gi<u>f</u>t usw. und

(49) flie<u>h</u>en ⟶ Flu<u>ch</u>t usw. nicht einmal "zumindest sehr ähnlich" (Wurzel, 239). Während nämlich bei den Verben unter (39) das stammauslautende konsonantische Segment andere distinktive Eigenschaften annimmt, also als Segment erhalten bleibt, verlieren die Verben unter (49) ihr stammauslautendes konsonantisches Segment, welches z.B. im Althochdeutschen sowohl im Verb (ahd. flio<u>h</u>an) als auch im Substantiv gleichzeitig vorzufinden war. Zweitens ist Wurzel zu vorschnell in der Annahme, dass bei <u>fliehen</u> usw. ein stammauslautender Konsonant in der Basisform vorhanden <u>ist</u>. Ganz richtig <u>war</u> er vorhanden, z.B. noch im Mittelhochdeutschen; im Neuhochdeutschen ist er es aber nicht mehr (Schirmunski, 364). An dieser Stelle ist der wichtige Hinweis zu machen, dass das Kind, das seine Muttersprache durch Aufbau einer Grammatik (= Modell der Kompetenz) erlernt, nichts von Etymologie und Lautveränderung wissen kann (King, 102; Wurzel, 207; Ladefoged, 48). Wer phonologische Regeln schreibt, sollte

beim Abfassen dieser Regeln nicht nur daran interessiert sein, verschiedene, auf irgendeine Weise scheinbar zusammengehörende Formen durch technische und formale Finessen vermeintlich kostensparend zu vereinen, sondern auch daran denken, dass diese Regeln u.a. psychologisches Geschehen und das Wissen des Sprechers wiederspiegeln sollten. Leider ist diese Forderung viel leichter aufgestellt als erfüllt (vgl. Ladefoged, 47 ff). Deshalb sind meine Gesichtspunkte und Lösungsvorschläge ebenfalls nur (aber hoffentlich mehr der Wirklichkeit entsprechende) Annahmen, die noch der Bestätigung durch das Experiment harren.

Das substantivbildende t-Suffix der Gruppen (39) und (49) ist heute ein unproduktives Suffix (Fleischer, 172 f), d.h. das Gegenwartsdeutsch enthält keine Substantivbildungsregel der Form

(1) Verb + t-Suffix ⟶ Nomen

mehr. Überhaupt ist ja das gesamte Ablautsystem heute nicht nur erstarrt, d.h. unproduktiv, sondern schon im Schwinden begriffen. Verschiedene, früher starke Verben, werden heute regelmässig flektiert, z.B. melken, hauen, schnauben. Nach Fleischer (S.172f) sind Paare wie <u>fahren - Fahrt</u>, <u>nähen - Naht</u>, ohne weiteres semantisch und morphologisch verbunden. Der Zusammenhang in Paaren wie <u>tragen - Tracht</u>, <u>ziehen - Zucht</u>, <u>fliehen - Flucht</u> usw. ist "etwas weniger deutlich, aber doch noch erkennbar". Ich möchte einschränkend hinzufügen: wohl vom Sprachforscher, aber keinesfalls von Otto Normalverbraucher erkennbar. "In Paaren wie <u>pflegen - Pflicht</u>, <u>biegen - Bucht</u>, <u>drehen - Draht</u> ist der semantische Zusammenhang verdunkelt, so dass man Substantiv und Verb auf <u>synchronischer</u> Ebene nicht mehr in einem Stammorphem vereinigen kann" (meine Unterstreichung). Auf die Situation des Spracherlernens durch das Kind bezogen, besagt dies, dass es z.B. <u>verzeihen</u>, <u>Verzicht</u>, und <u>bezichtigen</u> als drei unabhängige, ohne Beziehung zueinander stehende Wörter lernen muss. Das gleiche gilt entsprechend für z.B. mögen, mochte, Macht; bringen, brachte; nach, nahe. Auch wenn sie zu einem früheren Zeitpunkt der Sprachgeschichte eindeutig semantisch und phonologisch verbunden waren.

Die Paare under (49) und (50), die auf einer älteren Entwicklungsstufe aus einer zugrundeliegenden Form abgeleitet wurden, werden also heute beim Aufbau der Grammatik durch das Kind als separate Elemente ins Lexikon aufgenommen. Durch die historische Lautveränderung bei den Verben unter (49) sowie den Paaren unter (50), wobei das [x] intervokalisch ausfiel, entstand eine Umstrukturierung des Lexikons ("restructuring", King, 79 ff): Die Basisformen (Infinitivformen) dieser Verben wurden verkürzt zu /flī/ usw., die dazugehörigen Substantivformen, die das [x] vor $\#\#$t erhalten hatten, wurden neu aufgenommen, während sie auf der vorhergehenden Sprachstufe mit Hilfe der Substantivbildungsregel (1) abgeleitet wurden. Entsprechendes gilt für die Gruppe (50). Dass die Stammformen der drei Verben der Gruppe (51) ebenfalls als verschiedene Formative im Lexikon anzusetzen sind, dürfte sich nach dem bisher Gesagten von selbst ergeben. Selbst Wurzel (S.63) spricht von diesen Verben als "vollständig unregelmässige Verben".

Auch aus einem anderen Grund lehne ich die Existenz des (/ɣ/) ab. Das zugrundeliegende Konsonantensystem, das /ɣ/ enthält, ist asymmetrisch. [ɣ] taucht nur in einigen deutschen Dialekten auf der phonetischen Ebene als Realisation des zugrundeliegenden /g/ auf, z.B. Tage [tāɣə]. Ein phonologisches System mit /ɣ/ als einzigem stimmhaften spirantischen Segment ist nach unseren gegenwärtigen Kenntnissen unnatürlicher und darüberhinaus kostspieliger als eines ohne. Noch dazu macht die Tatsache, dass dieses Segment - verglichen mit dem gesamten Inhalt des Lexikones - nur bei einer äusserst geringen Zahl von Wörtern auftreten sollte, die Annahme dieses /ɣ/ unhaltbar.

Ganz besonders ist sie dies angesichts der Forderung, dass die Regeln der Grammatik das Wissen des Sprechers ausdrücken müssen (Ladefoged, 47 ff). Unter der rein theoretischen Annahme, dass Wurzels Regel (K 10') (S.241) dieser Forderung entspräche, muss der Sprecher des heutigen Deutsch die Substantivbildungsregel (1) kennen, hat er im Lexikon die Repräsentationen /fliɣ/ usw., benutzt die Regel (K 10') für die Prozesse unter (39) und (49) und verwendet schliesslich eine Regel (K 11d), um die verbleibenden /ɣ/-Elemente in den Verbformen zu eliminieren. Die Annahme, dass die Substantivbildungsregel (1) aktiver Bestandteil der Grammatik des Gegenwartsdeutsch sei, wurde schon weiter oben zurückgewiesen. Ebensowenig haben für den Sprecher des heutigen Deutsch die Prozesse unter (39) und (49) weder Ähnlichkeit miteinander noch sind sie in seiner Grammatik enthalten. Dieser angenommene Vorgang, bedingt durch die Annahme eines /ɣ/, scheint psychologisch völlig unrealistisch zu sein. Durch die Wahl des zugrundeliegenden /ɣ/, die einfach das Ergebnis der Suche nach einem Konsonanten war, der "hier ... funktioniert" (Wurzel, 239), findet Wurzel einen Trick, mehrere Erscheinungen, nämlich (39), (49) und angedeuteterweise auch (50) und (51) zu generalisieren, die wirklich nichts miteinander zu tun haben, weder geschichtlich, der Lautenwicklung gemäss, noch gegenwärtig, den Verhältnissen im Gegenwartsdeutsch entsprechend. Und diesen Kunstgriff nur, um sich keine scheinbar möglichen Generalisierungen entgehen zu lassen, um zu vereinfachen, um Kosten zu sparen. Mit Ladefoged und Ohala (19 ff) ist solch ein Verfahren zumindest abzulehnen, da es sich gezeigt hat, dass die von Chomsky und Halle angenommene Bewertungsprozedur, die für Wurzel oberstes Gesetz zu sein scheint, mangelhaft, wenn nicht ganz falsch ist.

Um zu demonstrieren, dass mit formalistischen Kniffen alles möglich ist, setze ich bei den 5 Verben unter (49) und den 3 Paaren unter (50) ein /x/ als stammauslautendes konsonantisches Segment an, z.B. /flix̲/, /xox̲/.

Diese Formen weisen, was das unterstrichene /x/ betrifft, die Verhältnisse älterer Sprachstufen auf, z.B. des Althochdeutschen. Durch Substantivbildungsregeln und Komparationsregeln werden die entsprechenden Nominalisierungen bzw. Steigerungsformen abgeleitet. Jetzt braucht nur noch das /x/ bei allen Verbformen und bei den Adjektiven dann eliminiert zu werden, wenn es in bestimmten Umgebungen auftritt. Bei den Verbformen fällt das /x/, mit Ausnahme der Formen des Präteritum, wie bei den Adjektiven intervokalisch, d.h. vor Morphemgrenze, aus. Es gilt also z.B.:

$$\text{(Regel 87)} \qquad x \longrightarrow \emptyset \Big/ \left\{ \begin{array}{l} \_\_\_\_\_\_ + \\ \left[\begin{array}{c} \_\_\_\_\_\_ \\ +\text{Prät} \end{array}\right] \end{array} \right\}$$

Ganz richtig liefert Regel 87 die erwarteten Formen, z.B. [fliən] aus /flix+n/, [näə] aus /nax+e/ usw.

Allerdings stellt ziehen - zog - gezogen auf jeden Fall eine Ausnahme dar, weil das g des Präteritum und des Partizip Perfekt nicht voraussagbar ist. Die Form zog muss also, wie man es nun drehen und wenden will, ins Lexikon aufgenommen werden.

Wenn Formative wie fluchen, seichen (Wurzel, 239) oder Buch, bei denen das /x/ in allen Formen erhalten bleibt, ebenfalls mit /x/ repräsentiert werden, müssen die Paare unter (49) und (50) durchgehend mit dem Merkmal [+Regel 87] versehen werden. Demjenigen aber, der diese Massnahme unschön findet oder dem sie zu viele Merkmale kostet, bietet sich sofort ein Ausweg an: Entsprechend der lautgeschichtlichen Entwicklung (Teil der hochdeutschen Lautverschiebung) können die gegenwärtigen [x]-Segmente auf der phonetischen Ebene der Wörter fluchen, seichen, Buch usw. als zugrundeliegende /k/-Segmente repräsentiert werden, die dann mit Hilfe einer einfachen Regel in ihre entsprechenden dauernden Segmente in der genau definierten Umgebung verwandelt werden:

$$k \rightarrow x \;/\; \_\_\_\_\_ \;\#$$

Diese Darstellung bietet einen einzigen Vorteil: sie benötigt kein zugrundeliegendes /ɣ/. Alles andere aber sind Finten, die sich angesichts der Forderungen nach Erklärungsrelevanz und Realität der Grammatik selbst unmöglich machen. Dazu kommt noch die Feststellung, dass das Kind und der nicht vorbelastete Sprachteilnehmer keine lautgeschichtliche Kenntnisse besitzen. Regeln, die Lautveränderungen in der Sprachentwicklung, und nur diese, wiederspiegeln, haben wohl ihre Daseinsberechtigung in einer lautgeschichtlichen Beschreibung, keineswegs aber in der Grammatik einer Sprache. Interessanterweise argumentiert Wurzel auf S.207 ganz im oben dargelegten Sinne und damit gegen seine eigene Behandlung der aktuellen Wortgruppen.

Somit sprechen verschiedene Argumente gegen die Annahme, dass das Deutsche einen stimmhaften velaren Spiranten im zugrundeliegenden Konsonantensystem hat. Diese neue Erkenntnis erfordert einige Berichtigungen in Wurzels Studien zur deutschen Lautstruktur (Alles andere soll als richtig stehen bleiben):

Die Regel (K 10) vereinfacht sich um das Merkmal [-sth] (S.238).

(K 10) <u>Wechsel vor t</u>

$$[+obstr] \rightarrow [+dnd] \;/\; \_\_\_\_\_ \;\# \begin{bmatrix} +obstr \\ +kor \end{bmatrix}$$

Teil (d) der Regel (K 11) <u>Velare Spiranten</u>, wonach verbleibende /ɣ/-Segmente getilgt werden, z.B. in /fliɣ+n/, entfällt.

Die Reihenfolge der Anwendung der Regeln ändert sich (Wurzel, 288 und Argument S.239, 2.Absatz). Die Ableitung für die Substantive der Gruppe

(39) geht so vor sich:

1. Ablaut des Verbes, z.B. grab ⟶ grub
2. Nominalisierung durch t-Suffix: grub # t
3. Auslautverhärtung (Stimmlosigkeit K 13): grup # t
4. Wechsel vor t (K 10): gruf # t

Die Regel (P 4) Gespanntheit der Ablautvokale (Wurzel, 75) kann bestehen bleiben. Die Stammformen von sehen und geschehen werden durch Teilregel (a) richtig abgeleitet; die von ziehen müssen wegen des nicht voraussagbaren g im Lexikon aufgenommen sein; die von fliehen und (ver)zeihen ebenfalls, da für diese beiden +PP=Prät, sie aber kein stammauslautendes Segment enthalten.

Damit ändert sich die Regelfolge insgesamt, verursacht durch das Streichen von zugrundeliegendem /z/ und /ɣ/, wie bereits auf S. 13 dargestellt.

Schliesslich ist noch zur Behandlung des Wortes Häher Stellung zu nehmen (Wurzel, 244). Ebenso wie Wurzel das Einschieben einer Morphemgrenze als technischen Trick ablehnt, also /xǟ+r/, muss seine Lösung, die "plausible Basisform" mit morpheminlautendem /ɣ/, also /xǟɣr/, das dann später wieder eliminiert wird, als zumindest ebenso übler technischer Trick verworfen werden. Allein schon die Tatsache, dass der Name dieses Vogels lautnachmenden Ursprungs ist, macht deutlich, dass dieses Wort, wie alle onomatopoetischen Bildungen, eine Sonderstellung im Wortschatz einnimmt. Im Hinblick auf das oben Ausgeführte bleibt also als einzig plausible und realistische Lösung übrig, die Form /xǟer/ Häher vs. /xēr/ Heer im Lexikon zu repräsentieren.

Zu guter Letzt noch eine Preisfrage: Was fangen wir mit dem etymologisch zusammengehörenden Paar rauh – Rauchwerk (=Pelzwerk) an? Die beste Lösung, die mir möglichst bald zukommen sollte, wird durch Veröffentlichung an anderer Stelle ausgezeichnet!

Anmerkung: Aus technischen Gründen mussten alle ß, bis auf wenige, für die Argumentation unentbehrliche Fälle, durch ss ersetzt werden. Trotz dieses Schönheitsfehlers ändert sich nichts am Inhalt der Besprechung.

ARE THERE UNDERLYING VOICED FRICATIVES IN GERMAN?

Summary in English

In his "Studien zur deutschen Lautstruktur" Wurzel proposes underlying /z/ and, only for a few cases, /ɣ/. There is no underlying /v/ in German. I try to show and argue that there are no underlying voiced fricatives in German at all.

The phone [z] always appears (i) when word-initial and prevocalic, and (ii) between voiced segments. But [s] can also appear between diphthong or long vowel and a following vowel. The S-Voicing Rule is given on p. 12 and inserted into Wurzel's order of rule application on p. 13. From the historical point of view, German has Germanic s (simple s) and ss which developed from Germanic t during the High German Soundshift. Both kinds of s are reflected in the orthography: ss (or ß) always denotes [s], s represents either [s] or [z] depending on the context. I suggest the following underlying forms: (i) /ss/ for the MHG ss or ß <OHG t, giving [s] by the Cluster Simplification Rule (p.11), already incorporated in Wurzel's "Studien", and (ii) /s/ for the Germanic simple s, giving either [z] or [s] by the S-Voicing Rule.

To postulate /ɣ/ is just a trick, to save costs and to generalize where generalisations are not possible. Wurzel uses /ɣ/ to account for a pseudo-alternation in a few pairs of words (p.13). In present German these words are rather isolated and to satisfy the claim of psychological reality they should be represented in the lexicon independently of each other, viz. without /ɣ/.

Literaturverzeichnis

Bierwisch M. 1966. Regeln für die Intonation deutscher Sätze. Studia Grammatica VII

Chomsky N. und Halle M. 1968. The Sound Pattern of English

Fleischer W. 1969. Wortbildung der deutschen Gegenwartssprache

King R.D. 1969. Historical Linguistics and Generative Grammar

Kiparsky P. 1966. Über den deutschen Akzent. Studia Grammatica VII

Kitzing K. 1967. Några malmöelevers uttal av tyska spiranter och affrikator. Pedagogisk-psykologiska problem. Nr. 49. Malmö, Schweden

Ladefoged P. 1971. The Limits of Phonology. Form and Substance, herausgegeben von L.L.Hammerich, R.Jakobson und E.Zwirner. Kopenhagen

Ohala J. 1971. The Role of Physiological and Acoustic Models in Explaining the Direction of Sound Change. Vervielf.

Ross J.R. 1967. Der Ablaut bei den deutschen starken Verben. Studia Grammatica VI

Schirmunski V.M. 1962. Deutsche Mundartkunde. Berlin

Stanley R. 1967. Redundancy Rules in Phonology. Language 43:393-436

Wurzel W.U. 1970. Studien zur deutschen Lautstruktur. Studia Grammatica VIII

Zimmer K. 1969. Psychological correlates of some Turkish morpheme structure conditions. Language 45:309-321

LARYNGEAL BOUNDARY SIGNALS

Eva Gårding

In this report we shall study how a variation in morpheme boundary location may affect the behaviour of some laryngeal muscles (vocalis m, cricothyroid m, and sternohyoid m). The activity of the muscles was recorded in an EMG investigation carried out at the Research Institute of Logopedics and Phoniatrics in Tokyo during the fall of 1969.[1] The present data are derived from a speech sample which was composed chiefly to investigate how the selected muscles participate in the production of Swedish word tones. A preliminary report of the word tone data (Gårding et al., 1970) was published in the Annual Bulletin No. 4. The same bulletin contains reports on a number of other experiments all of which aim at exploring the functioning of laryngeal muscles during speech (Hirose et al., Simada and Hirose, Ohala and Hirose, 1970).

MORPHEME BOUNDARIES AND SYLLABLE BOUNDARIES

Acoustic and perceptual aspects of morpheme and syllable boundaries were studied rather extensively in Swedish material (Gårding, 1967).

A spectrographic analysis of pairs of sequences that differ in the location of a morpheme boundary (e.g. tåg-ångare, "train-ferry", and tå-gångare, "toe-walker") showed that the acoustic segments around the boundary depend on the structure of the underlying morphemes, the stress pattern of

---

1 I wish to thank my collaborators at the Institute of Logopedics and Phoniatrics for their invaluable aid. I am particularly indebted to Doctors O. Fujimura, H. Hirose, and Z. Simada. I have also profited from some comments from Doctor M. Sawashima.

the phrase and the rate of utterance of the speaker. Between stressed syllables the speaker can time and arrange his speech gestures in different ways. The phonetic result was called internal juncture, a marked boundary. By means of perceptual tests it could be shown that listeners react to these differences particulary if they are accompanied by a change in the feature composition (aspiration, glottal stop etc.) of the segments involved. The spectrographic study also brought out some common characteristics by which speakers realize morpheme boundaries in varying phonetic environments. The location of the boundary can in most cases be related to an interval of low intensity in the spectrogram. From a study of the formant movements it could also be inferred that the speech organs slow down and move toward a neutral position in connection with the marked boundary. When the second morpheme starts with a vowel, the airstream is checked by a glottal closure or constriction (various manifestations of glottal stops appear in the spectrograms). In a sequence in which the second morpheme has a consonantal beginning there is no spectral indication of a glottal closure. However, the initial consonant closure or constriction is typically prolonged by a time interval comparable to that of the glottal closure.

With an increased rate of utterance speakers change their articulation in a uniform manner. One of the stressed syllables is reduced, the prolonged closure or constriction interval disappears and the intervocalic consonants are rearranged in such a way that as many as possible become syllable initial. This was interpreted as an adjustment on the part of the speaker to the general syllabification patterns (the unmarked boundaries) of the language. The process can be seen as the result of a shift to a simpler production program. In an informal test speakers appeared to have surprisingly similar notions of the syllable boundary locations for which rules could be set up. Perceptual tests showed that when boundaries followed these rules the morpheme division made by listeners tended to become haphazard. For example, a syllabification like tå-gångare in rapid speech could be interpreted as meaning both train-ferry and toe-walker.

MATERIAL

The present material was composed in such a way that it will be possible to study the effect of moving the morpheme boundary in a sequence from ...V + CV... to ...VC + V... and the effect of replacing a morpheme boundary ...V + CV... by a syllable boundary ...V - CV...

Test words

Orthographic representation with hyphens to show constituents	Modified IPA transcription
(1) mån-år or må-når	[ˇmo:no:r]
(2) må-når	[ˇmo:ˋno:r]
(3) mån-år	[ˇmo:nˋo:r]
(4) må-når	[ˇmo:'no:r]
(5) mån-år	[ˇmo:n'o:r]
(6) tå-gångare or tåg-ångare	[ˇto:gɔnaˋrɛ]
(7) tå-gångare	[ˇto:ˋgɔnarɛ]
(8) tåg-ångare	[ˇto:gˋɔnarɛ]

The transcription follows the IPA principles according to which /'/ indicates strong stress with the acute (simple) tone and /ˇ/ marks the stressed syllables of words with the grave (compound) tone. When two or more syllables follow the stress, the syllable bearing the second element of the compound tone is marked by /ˋ/. In the emphatic utterances (4) and (5), underlined in the orthographic representation, /ˋ/ is replaced by strong stress /'/.

The test words are compound nouns made up of Swedish morphemes. Number (4) is contrived and specially designed for the experiment. Conforming to the rules for compound formation, the stressed syllable of the first element carries the grave accent /ˇ/.

Items (1) and (6) are meant to exemplify a pronunciation which conforms to the general syllabification rules (in this case ...V: - C...). It has a reduced stress on the second element of the compound. This syllabification will be used in relaxed speech when the compound has a high degree of expectancy and the speaker does not insist on the meaning of the individual morphemes.

In sentences (4) and (5) the test words were given contrastive stress. In the context used here the speakers used a stress pattern which adds prominence chiefly to the second element of the compound.

The test word manär is composed of segments that minimize the effect of articulation on laryngeal muscles and is expected to reveal the effect of prosodic gestures (tone, stress). In the test word tågångare on the other hand, the t- and g-gestures are expected to influence laryngeal activity.

All of the test words were put in the frame dɛ.va - han sɑ (it was - he said) and each test sentence was uttered 15 times in a series with a short pause after each item.

Subjects

The subjects were L (male) speaking Standard Central Swedish and E (female) speaking a Skåne dialect. The different pitch contours by which the grave accent is manifested are typical of the dialects. (Compare Figures 5, 6, 7, 8.) Figures 1 and 2 show some spectrographic examples of the test words.

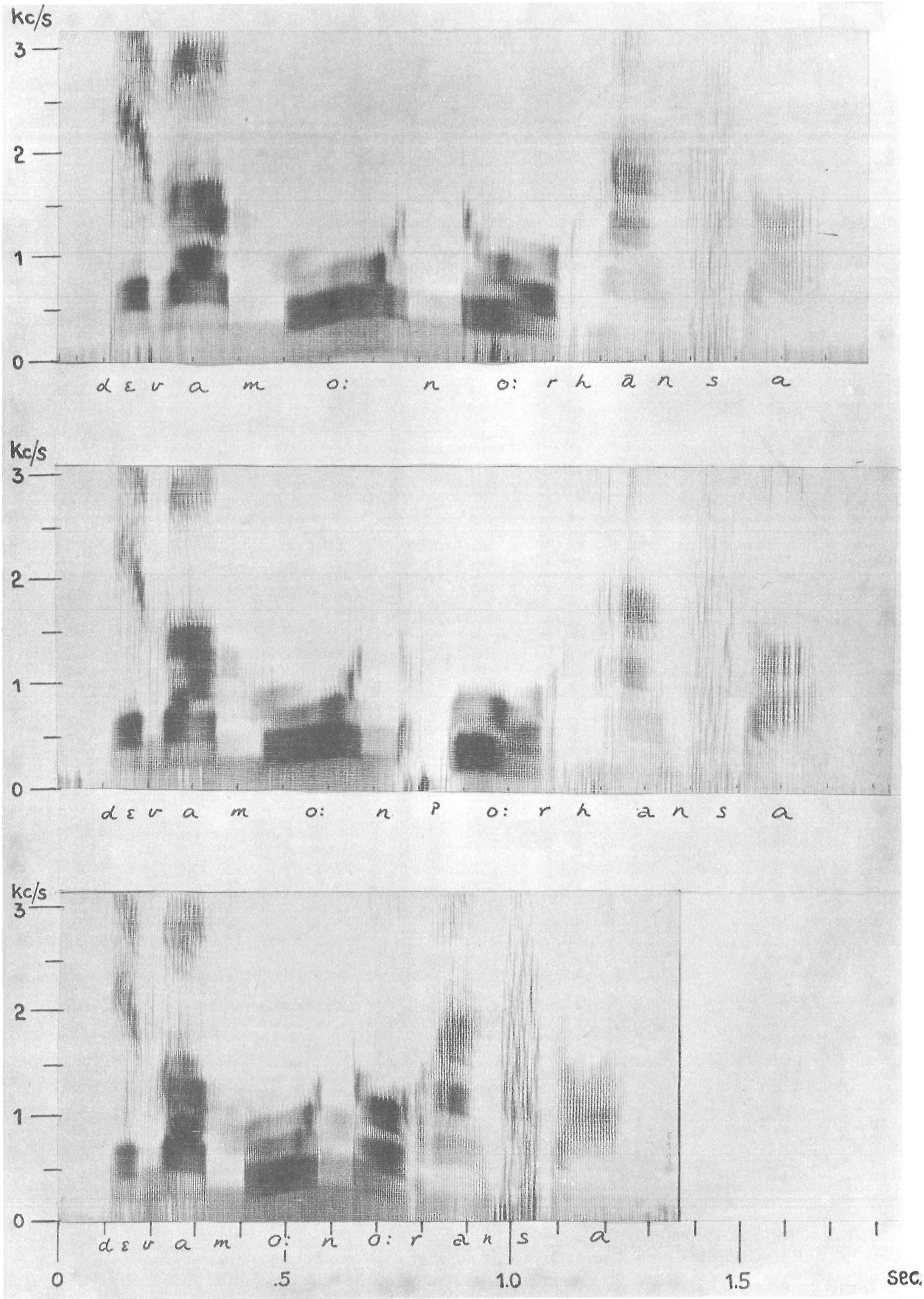

Figure 1. Spectrographic examples of the sequence <u>månår</u> pronounced with different boundaries. From top to bottom, [ˇmo:ˋno:r], [ˇmo:nˋo:r], [ˇmo:no:r]. Frame, <u>dε va</u> - <u>han sa</u>. Speaker E.

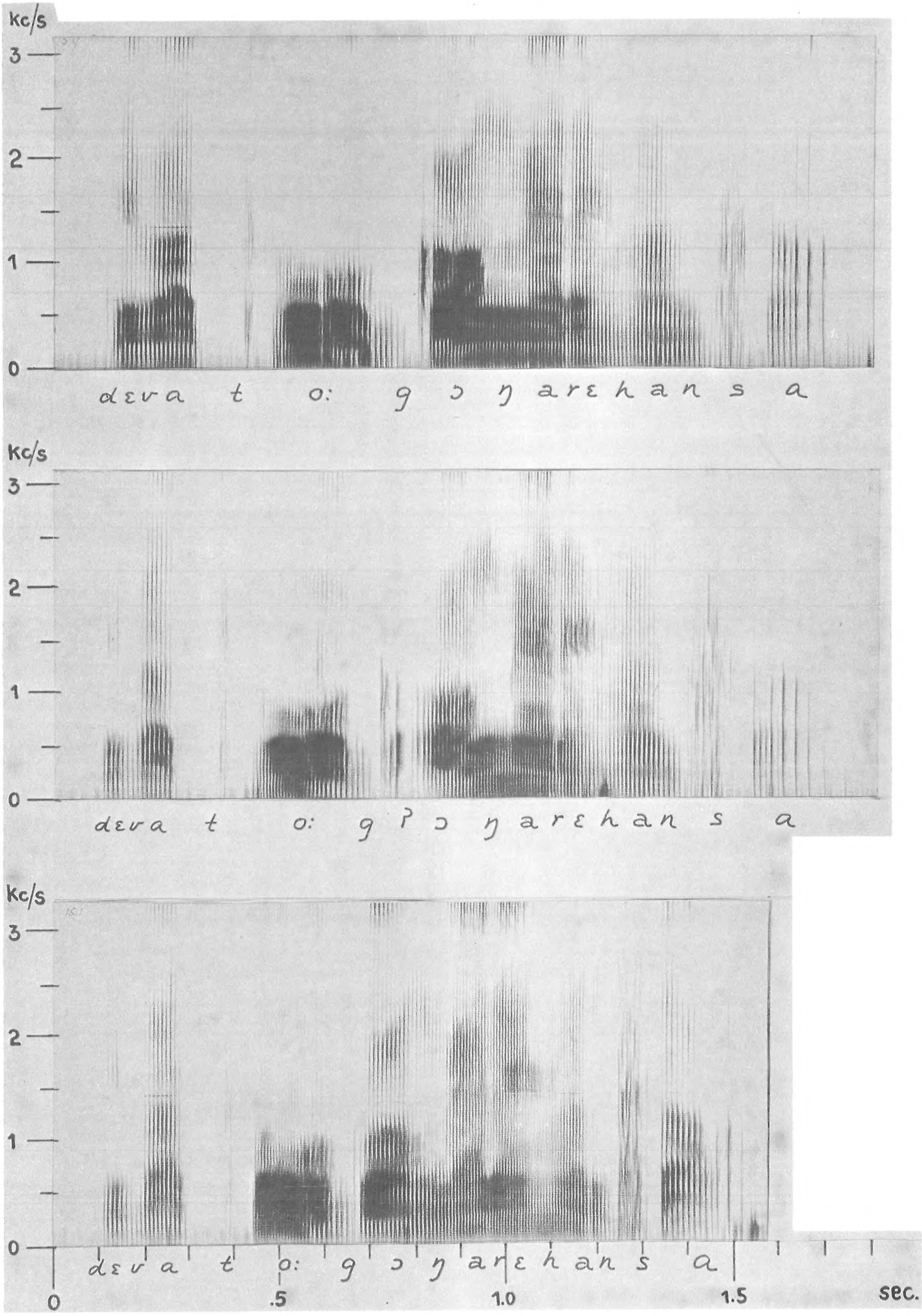

Figure 2. Spectrographic examples of the sequence tågångare pronounced with different boundaries. From top to bottom, [ˇto:ˋgɔŋarɛ], [ˇto:gɔ̀ŋarɛ], [ˇto:gɔŋarɛ]. Frame dɛ va – han sa. Speaker L.

EXPERIMENTAL PROCEDURES

The EMG data were obtained by means of double-ended hooked wire electrodes which were inserted through the skin and other tissues of the neck. (For a full description of the technique see Hirose et al., 1970.) Figure 3 shows the pertinent muscles.

The amplified EMG signals and the speech signal which had been recorded simultaneously on magnetic tape were fed to a PDP-9 computer through an AD converter for data processing. In this process the EMG signals were sampled every 250 microseconds and the values were converted into 6-bit levels. The digitized values of these samples were then averaged over a running window with a range of 10 msec. Out of 15 utterances of each test sentence, 10 were selected by auditory analysis and processed in this way. The resulting records of each set of 10 were then lined up in time with respect to some easily identified speech event and superimposed giving the final record as shown in the figures. Different choices of line-up points proved to give very slight variations in the pictures. Figure 4 shows examples of our averaged and processed data.

EXPECTED EMG CORRELATES

From what we know of internal juncture in Swedish and the activities of the muscles under investigation we should expect the following EMG correlates:

1. Glottal stop

Vocalis m. There should be vocalis activity during the glottal stop in VC + V sequences. The vocalis muscle has been shown to be active for glottal stops (Faaborg-Andersen 1957, Ohala 1970, Hirose et al. 1970, Gårding et al.

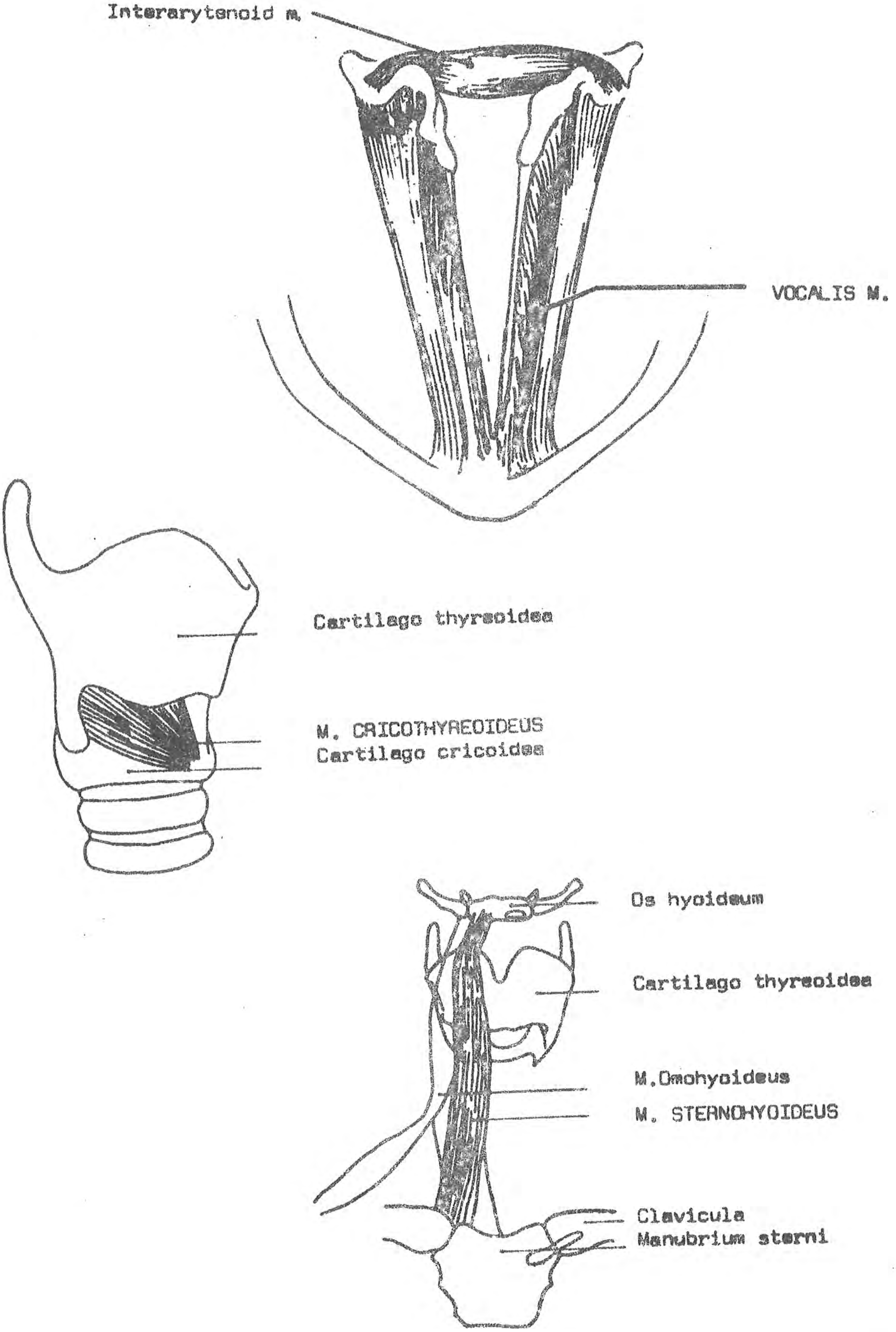

Figure 3. The target muscles. (From Ohala 1970, Hirano and Ohala 1969, and Kamijo 1966.)

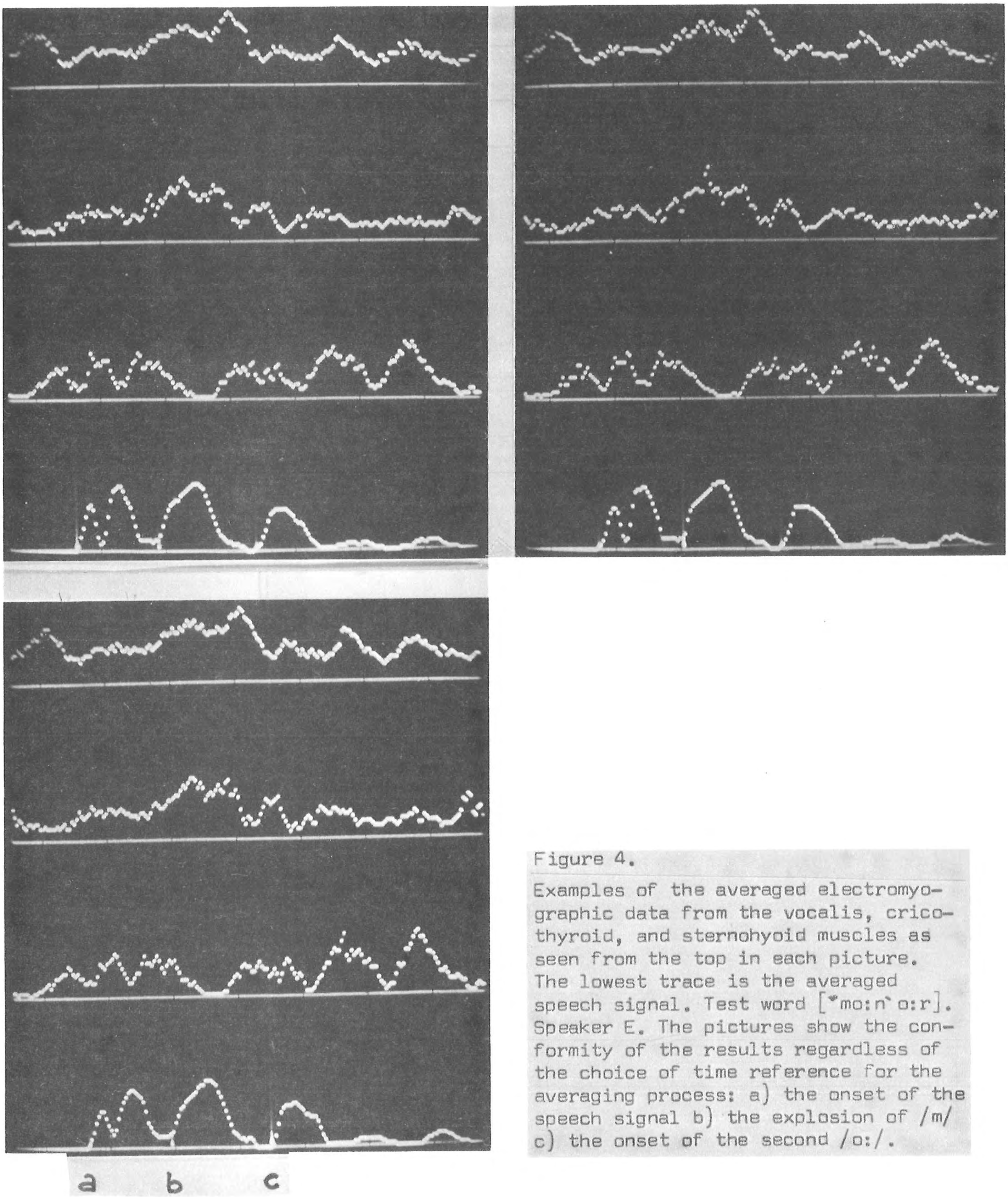

Figure 4.

Examples of the averaged electromyographic data from the vocalis, cricothyroid, and sternohyoid muscles as seen from the top in each picture. The lowest trace is the averaged speech signal. Test word [ˈmo:nˋo:r]. Speaker E. The pictures show the conformity of the results regardless of the choice of time reference for the averaging process: a) the onset of the speech signal b) the explosion of /m/ c) the onset of the second /o:/.

1970). Figure 5 shows for comparison conscious glottal stop gestures described in our earlier report. Notice that the cricothyroid activity seems to be suppressed at the time when the vocalis muscle has a peak of activity in connection with the glottal stops.

2. Timing

Sternohyoid m. According to spectrographic data (Gårding 1967) intervocalic C comes earlier in a VC + V sequence as compared to a V + CV sequence. When C = g, as in tågångare the C gesture is associated with tongue retraction which is known to involve the sternohyoid muscle (Ohala and Hirose 1970). Hence an earlier burst of activity is expected in tåg+ångare as compared to tå+gångare.

Cricothyroid m. In the previously mentioned EMG investigations of laryngeal muscles, the cricothyroid muscle appeared to correlate positively with the major movements of the pitch curve. Since the pitch curve is also influenced by variations in oral pressure caused by the articulation we can expect the cricothyroid record to represent a "cleaner" prosodic signal.

3. Stress

Reduced vocalis-cricothyroid activity is expected when stress is reduced as in the shift from a marked morpheme boundary to a syllable boundary. The vocalis and cricothyroid muscles have been shown to cooperate for pitch rising in stressed syllables. (See the works, cited above, and Sawashima et al. 1969.)

Contrastive stress is expected to enhance the difference in laryngeal activity associated with a shift of boundary. Contrastive stress was found to increase the EMG signal amplitude associated with consonant phonemes by 10-20 percent (Harris et al. 1968). A change of lexical stress had no such effect.

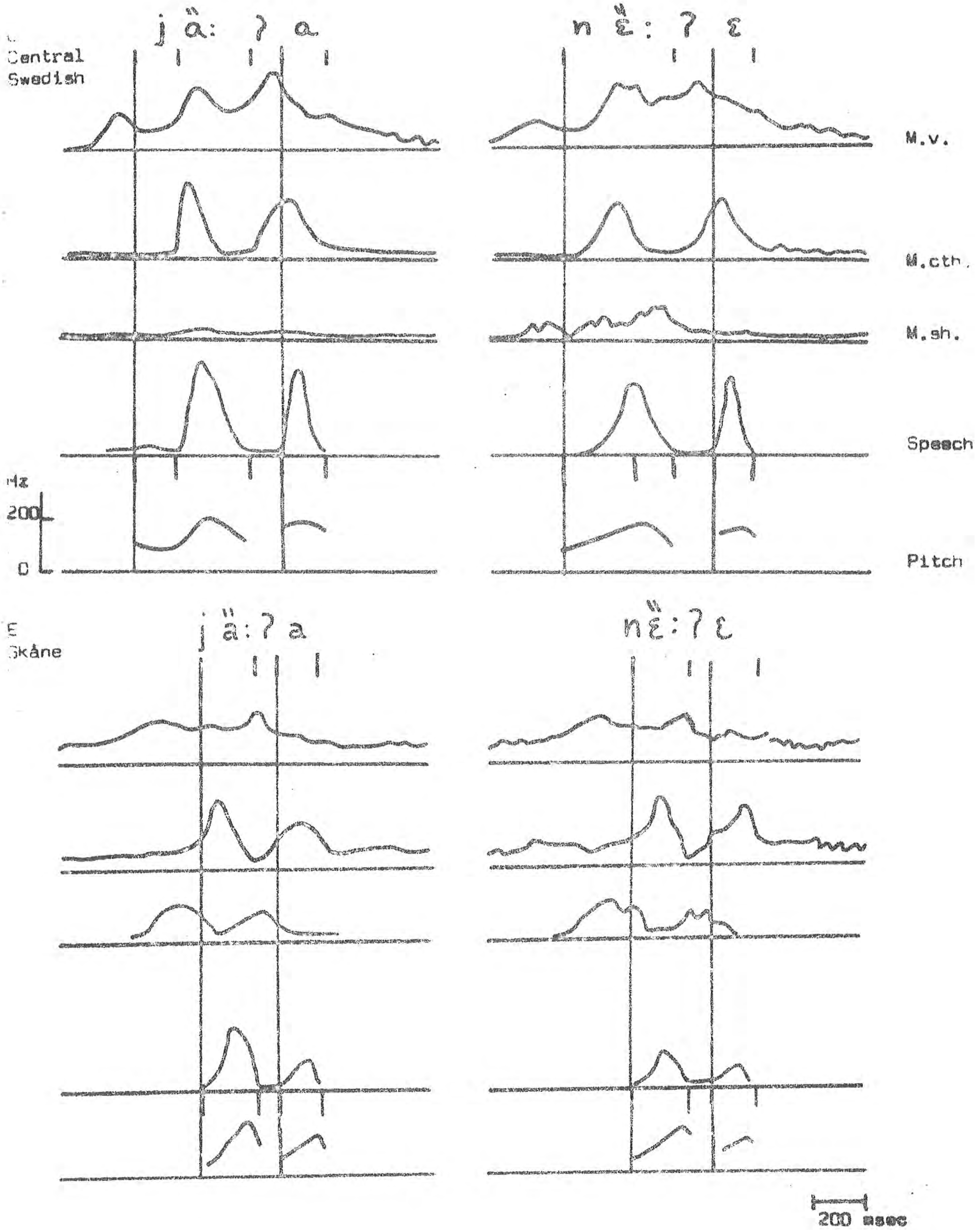

Figure 5. Data from expressive utterances with deliberately produced glottal stops. Speaker L above and E below. From Gårding et al. 1970

RESULTS AND DISCUSSION

Figures 6–8 show some results of the experiment. Each of the EMG-curves represents an average of 10 utterances. They have been derived from the vocalis, the cricothyroid and the sternohyoid muscles. The fundamental frequency curve shown as the lowest trace in the figure is the hand-made average of three representative utterances.

Figures 6–7 are derived from Speaker E, Figure 8 from Speaker L. The first vertical line represents the beginning of the speech signal, the second line connects the reference points used for the summation process, i.e., for the test word <u>månår</u> the release of [m], and for <u>tågångare</u> the release of [t].

In the following we shall comment mainly on observations concerning boundary problems. For EMG correlates to other speech gestures in the material see Gårding et al. 1970, Gårding 1970, and forthcoming.

<u>Effects of a morpheme boundary shift</u>

<u>Glottal stop</u>. With a shift of boundary from ...V + CV... to ...VC + V... we notice: A gap in the broad band spectrogram before the postjunctural, initial vowel which is preceded by a schwa segment with relatively slow vocal cord vibration (Figure 1 a and b). The comparable spectrogram in Figure 2 b [ˇto:gˋɔngarɛ], derived from a test sentence with the same placement of juncture uttered by Speaker L, has an absolute gap of much shorter duration. However, this gap is followed by a segment caused by the creaky onset of the following vowel which because of its scant supply of spectral energy is similar to a gap. The duration of these two successive gaps in Speaker L's utterance is comparable to that of the single gap in Speaker E's. The perceptual effect must be similar in this respect: the segments

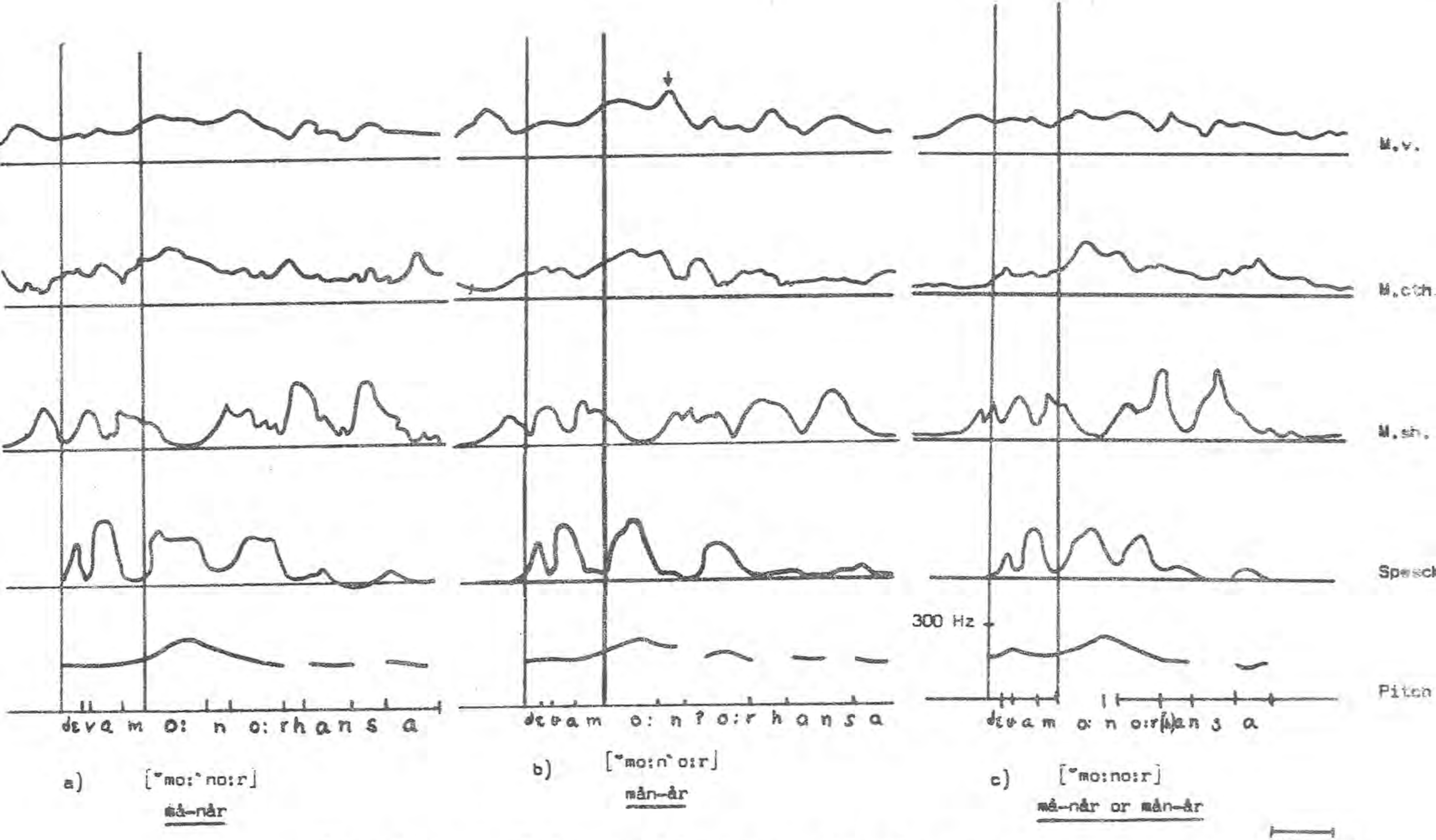

Figure 6. Speaker E. Sequence månår with three kinds of boundaries: a) [ˇmo:ˋno:r], b) [ˇmo:nˋo:r], c) [ˇmo:no:r]. The pitch curve is the average of three representative utterances.

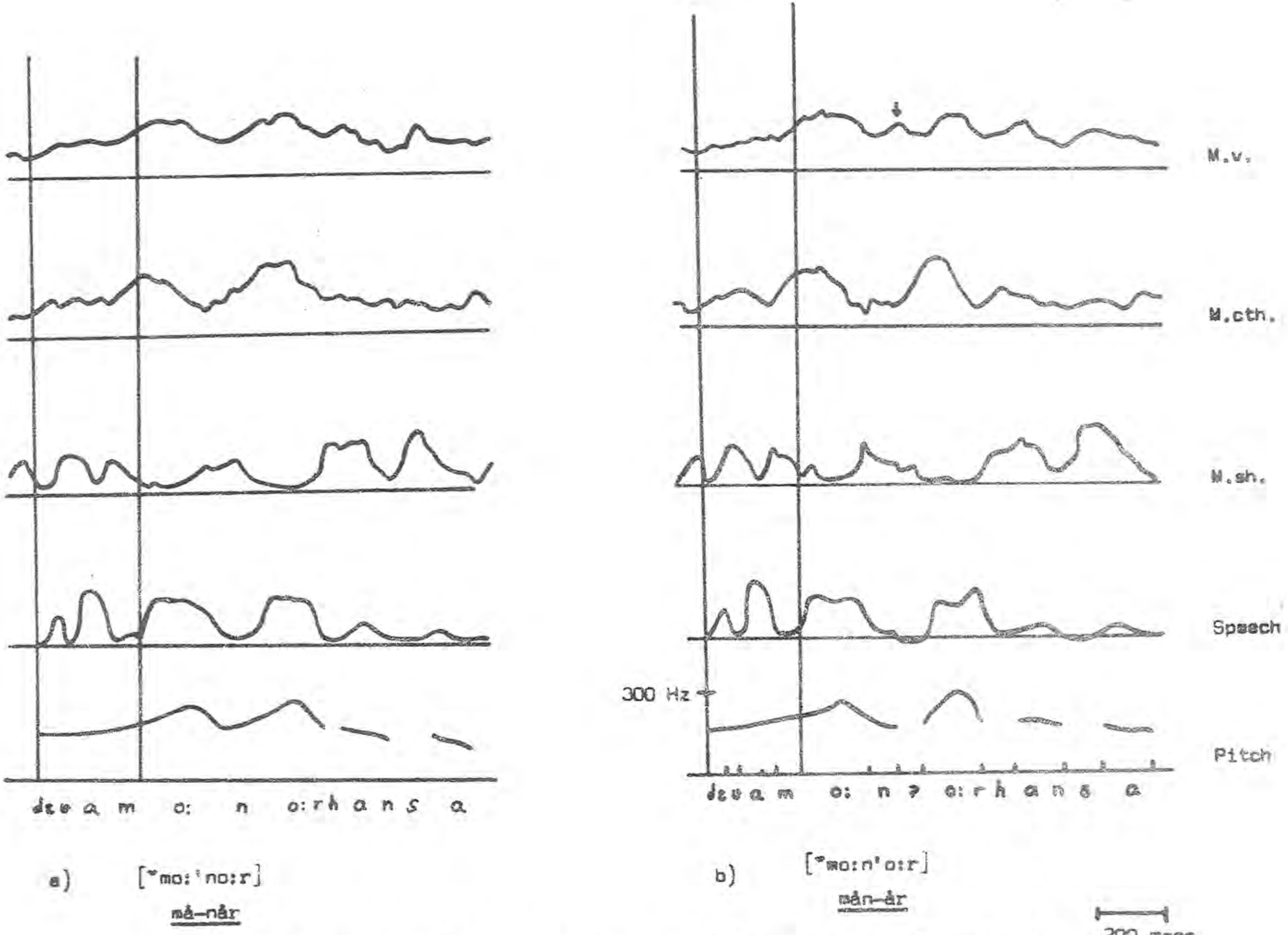

Figure 7. Speaker E. Contrastive stress. a) [°mo:'no:r], b) [°mo:n'o:r].

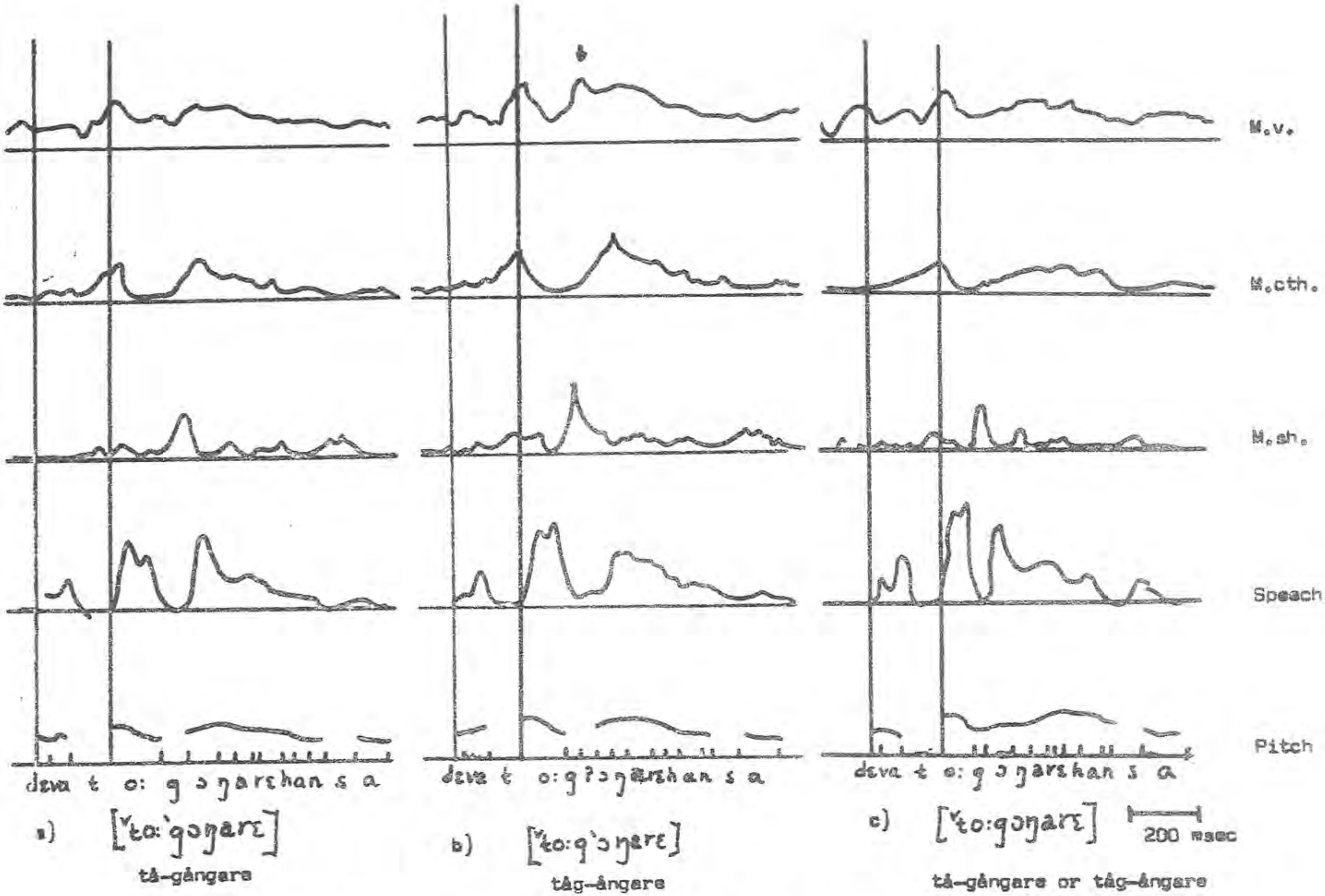

Figure 8. Speaker L. Sequence tågångare with three kinds of boundaries:
a) [ˇto:`goŋarɛ], b) [ˇto:g`ɔŋarɛ], c) [ˇto:goŋarɛ].

separate C+ from the following syllable kernel. It can be inferred from the spectrograms that Speaker E produces the boundary in a ...C + V... sequence by means of a glottal closure and a softly attacked vowel and that Speaker L makes a less complete closure which permits the vocal cords to vibrate slowly. For both speakers the vocalis muscle has an EMG peak which correlates to the acoustic gap and during this time the cricothyroid muscle is more or less passive. (An arrow points to these vocalis peaks in Figures 6, 7, and 8.)

This vocalis-cricothyroid interrelation is consistent in all the ...VC + V... sequences of our material. There is also a constant timing difference between the two speakers. For Speaker L (Figure 8 and also Figure 5) the vocalis peak associated with the glottal stop appears toward the end of the gap, for E it appears prior to the closure. It seems possible to associate E's earlier vocalis peak with the sudden and total glottal closure whereas L's peak may be connected with the creaky vowel onset.

The vocalis peaks observed in connection with the glottal stops may be interpreted as active laryngeal boundary gestures typical of the ...VC + V... sequence. Notice for cmparison the acoustically similar creaky utterance endings in sa (Figures 1 and 2) which are not connected with any vocalis activity.

We have seen that the cricothyroid and vocalis muscles, which have been found to cooperate for major pitch rises, have a different reaction pattern for glottal stops. Figure 9 (from B. Sonesson, Human vocal folds, to be published) illustrates the mechanical effects of contraction in these two muscles.* When the vocalis muscle contracts, the arytenoids will be pulled forward and the vocal folds will be shortened. This shortening effect can

---

* I am grateful to Doctor B. Sonesson for many explanations and discussions and for lending me the drawing appearing as Figure 10.

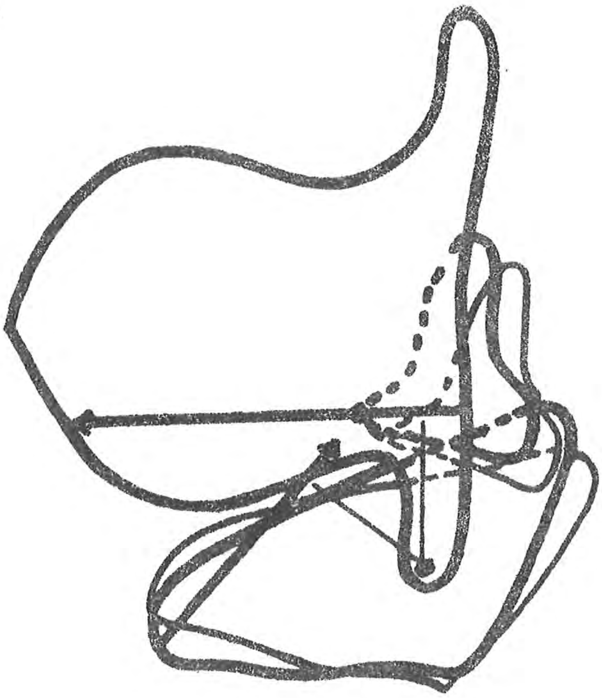

Figure 9. Mechanical aspects of vocalis and cricothyroid activity. The horizontal arrow shows the direction of the vocalis force. The slanted arrow indicates the force of the cricothyroid muscle. The straight lines drawn perpendicular to the arrows represent the levers from the cricothyroid joint. From Sonesson, Human vocal folds, to be published.

develop only if there is no counteracting force due to activity in a muscle working in the opposite direction as for instance the cricothyroid muscle. It has been found by visual inspection that the vocal folds are shortened and bulged for glottal stops. Hence it is natural that the cricothyroid muscle should not be activated for glottal stops. When the vocalis and cricothyroid muscles contract simultaneously the cricothyroid force prevents the vocalis muscle from being shortened. The vocalis activity will then produce the inner tension in the vocal fold needed to raise pitch.

Timing. With /g/ as the intervocalic consonant as in Figure 8, we can see the different timing of the intervocalic consonant gesture in the sternohyoid record. A comparison of [ˇto:gɔnarɛ] and [ˇto:gˋɔ narɛ] shows that the major sternohyoid peak, which is probably connected with the g-release, comes earlier when /g/ is syllable final both absolutely and in relation to the intensity maximum of the preceding vowel. In the given context it also has a faster buildup of activity. These findings are in agreement with the observations of the formant movements of the intervocalic consonant in the earlier study (Gårding 1967).

The cricothyroid activity connected with the pitch rise in the second syllable of the test sequence månår (Figures 6 and 7) is differently timed depending on the boundary location. When the intervocalic consonant is syllable initial, for instance in [ˇmo:'no:r], Figure 7 a, the cricothyroid curve starts rising prior to the /n/ segment. In [ˇmo:n'o:r] on the other hand the corresponding cricothyroid rise comes after /n/ obviously in connection with the initial vowel. These data suggest that the activation of the cricothyroid muscle for the pitch rise is tied to the beginning of the syllable independently of the syllable's articulatory characteristics.

The time interval during which the cricothyroid muscle is activated for the pitch movement of the first syllable is comparable in the two test words må-når and mån-år (Figure 7). This suggests that the duration of the prosodic signal is dependent on the degree of stress rather than the number of phonemes of the syllable. For the second syllable, however, the cricothyroid activation time is longer in når than in år. The cricothyroid activity starts rising prior to the /n/ in når whereas in år the rise starts in connection with the initial vowel . Earlier considerations of the vocalis-cricothyroid interaction make it natural to assume that in this case the glottal stop gesture interferes with the prosodic gesture and delays the pitch rise.

Stress. Figure 10 compares the test word må-når under normal (?) and contrastive stress. The two sets of curves have been superimposed with the beginning of the speech signal as their common reference line. It is obvious that the difference in muscular activity is mainly localized to the second syllable of the test word. The pitch curve derived from the contrastively stressed utterances has a prominent peak in this syllable which is obviously related to increased vocalis-cricothyroid activity. In spite of the fact that listeners still regard the first syllable as carrying the main stress (as shown in a test), there is no evidence that the EMG signal amplitudes for this syllable are influenced by the contrastive stress. We notice a difference in the duration of certain acoustic segments however. The consonantal occlusion is longer in the emphasized word both in the initial and intervocalic nasals. This suggests a different time program for the gestures in the two test words which is also evidenced by the sternohyoid record.

According to Fromkin (1966) a syllable initial /b/ has a stronger EMG signal (orbicularis oris) than a syllable final /b/. But Harris et al.

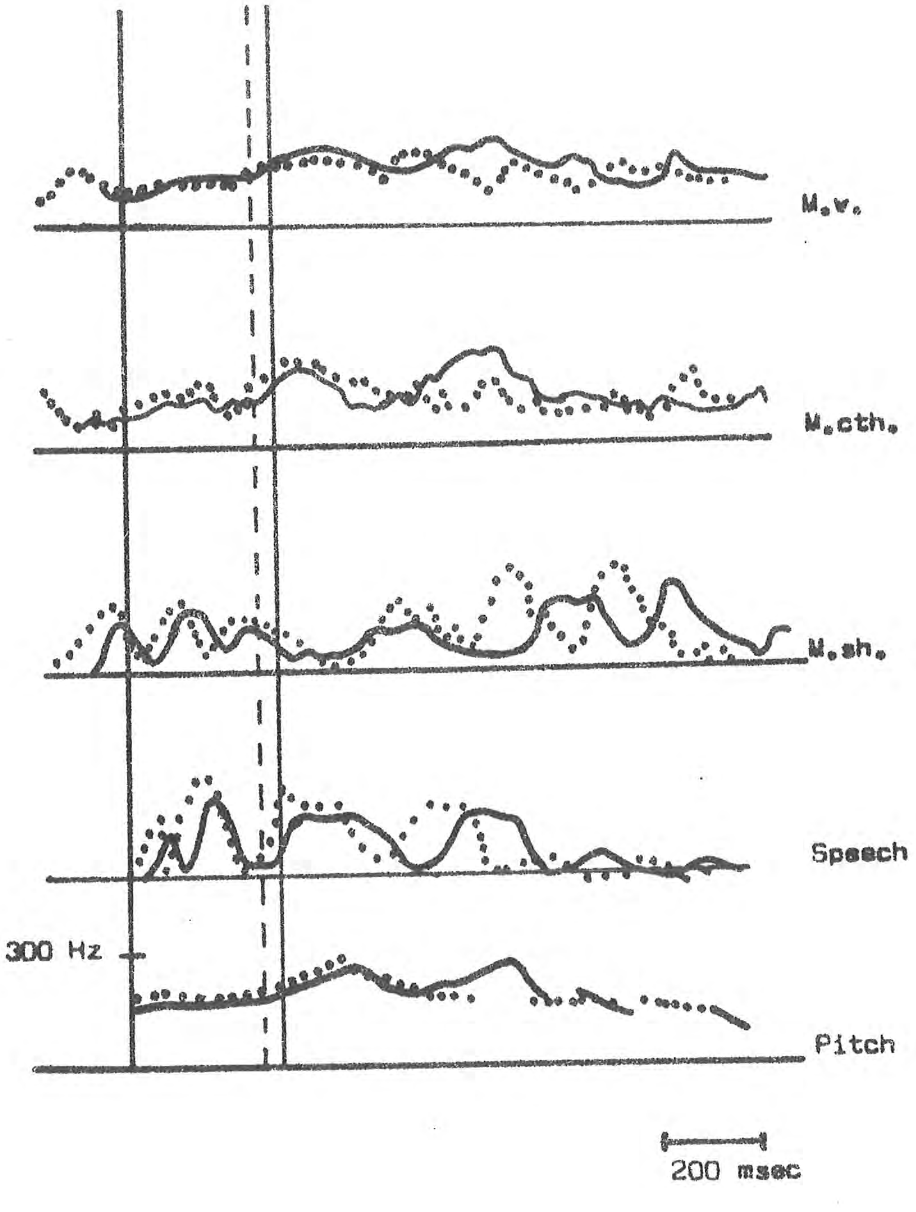

Figure 10. Speaker E. Sequence må-når with the same boundary location under normal (broken line) and contrastive (unbroken line) stress. The curves are superimposed with the onset of the following speech signals as their common time reference. The vertical lines are the time references for the averaging process.

(1965), found no significant difference in this respect. The present sternohyoid data suggest that the syllable final /g/ in 8 b requires a higher degree of muscle activity than the syllable initial /g/ in 8 a and c, an interpretation that has support in the spectrograms (Figure 2) which show a greater amount of acoustic energy in /g+/ as compared to /+g/. All these different results may perhaps be explained by the great stylistic variability of syllable final consonants. In this position a consonant may vary from weak to strong, and the strongest variant is often released with a following voiced or voiceless schwa element, which actually makes the final consonant comparable to a syllable initial one. In the context used here the speaker probably made a conscious effort to keep the consonant separated from the following syllable, hence the stronger burst of activity noted in the EMG signal.

To sum up our observations in Figures 6-8, the location of a morpheme boundary is traceable to the laryngeal muscles that have been the targets of this investigation. In a sequence ...VC + V... the boundary is controlled by the vocalis which contracts for the glottal stop after C+ and brings about a larger separation of the two syllable kernels. There is also some evidence in our EMG data that the activity of the cricothyroid muscle which regulates the observed pitch patterns is closely timed with the syllable.

Morpheme boundary and syllable boundary

Figures 8 a and 8 c illustrate the difference between a sequence pronounced with an unambiguous realization of a morpheme boundary [˘to:ˋgɔnarɛ] and the same sequence pronounced at a higher rate of utterance with an ordinary syllable boundary [˘to:gɔnarɛ]. The location of the boundary with reference to the segments involved is the same in the two sequences.

Stress. The test word [ˇto:ˋgɔnarɛ] has a higher degree of subjective stress in the second syllable than [ˇto:gɔnarɛ]. This higher stress does not produce a larger energy output but the pitch curves are different. In the first case the pitch starts falling earlier and the vocalis and cricothyroid muscles have higher and earlier peaks. In addition, the spectrograms show a longer occlusion and a longer open interval with stronger energy for the more stressed /g/.

Timing. Because of the longer duration of the segments pertaining to /g/ in [ˇto:ˋgɔnarɛ], the syllable kernels of the first two syllables are wider apart. The acoustic gap relates to a very small peak in the vocalis muscle and a passive phase in the cricothyroid. The combined behaviour of the two muscles actually looks lika a glottal stop gesture. – After the longer occlusion in [ˇto:ˋgɔnarɛ] there is of course a delay in the timing of the gestures for the rest of the utterance.

The two compared test words [ˇto:ˋgɔnarɛ] and [ˇto:gɔnarɛ] have as was already mentioned the same syllabic division. We notice that the major sternohyoid peak is similarly timed realtive to the intensity peak of the following vowel and has the same signal amplitude regardless of the difference in stress pattern. This relation shows that the sternohyoid peak is connected with the g-release. The timing of the peak in relation to the preceding vowel is different however. In [ˇto:gɔnarɛ] the peak has a fast rise, whereas in [ˇto:ˋgɔnarɛ] the rise is slow which may be indicative of a slower articulatory movement. The most conspicuous spectrographic difference in the two test words is a longer occlusion for /g/ in [ˇto:ˋgɔnarɛ].

The preceding observations may be summed up in the following way. With a shift of boundary from ...VC + V... to ...V + CV... to ...V – CV... the two syllables come closer. This process occurs at the expense of stress in the second syllable and at the expense of the duration of a glottal or oral closure or constriction. The glottal closure is most pronounced in the ...VC + V... sequence.

One other aspect of the same process can be expressed as follows. Two stressed syllables in succession seem to require a relatively long closure or constriction interval after the first syllable. During this interval the subglottal pressure needed for the second major stress is probably being restored.

Concluding remarks

On the basis of the spectrographic study (Gårding op.cit. p. 133 ff.) it was assumed that in sequences ...V_1 + CV_2... the articulatory program must be simpler than in sequences ...V_1C + V_2... Using Öhman's coarticulation model (Öhman 1967) the consonant in ...V + CV... could be regarded as superimposed on vocal tract shapes that gradually change from V_1 to V_2. In V_1C + V_2 on the other hand, C has resonances that indicate an intervening change in the vocal tract configuration from V_1 to schwa to V_2. In other words, C+ seemed to coarticulate mainly with schwa and +C with V_2. The glottal closure or constriction in connection with the initial postjunctural vowel in ...VC + V... also suggested a more complex innervation pattern.

The EMG data obtained from the vocalis and cricothyroid muscles in the present study are in agreement with the view that a ...VC + V... sequence requires a more complicated set of signals to the vocal apparatus than a ...V + CV... sequence. The glottal closure gesture in ...VC + V... calls for a differentiation in the activities of the two muscles which in the EMG records of the ...V + CV... sequences show a smooth cooperation.

References

Faaborg-Andersen K. 1957. Electromyographic investigation of intrinsic laryngeal muscles in humans. Acta Physiol. Scand. 41, Suppl. 140

Fromkin V.E. 1966. Neuromuscular specification of linguistic units. Language and Speech 9, 170-199

Gårding E. 1967. Internal juncture in Swedish. Travaux de l'Institut de Phonêtique de Lund VI. Lund

Gårding E. 1970. Word tones and larynx muscles. Working Papers, Phonetics Inst. Lund University 3

Gårding E., Fujimura O., and Hirose H. 1970. Laryngeal control of Swedish word tones. Annual Bulletin, Research Institute of Logopedics and Phoniatrics, University of Tokyo 4, 45-54

Harris K.S., Lysaught G.F., and Schvey M.M. 1965. Some aspects of the production of oral and nasal labial stops. Language and Speech 8, 135-147

Harris K.S., Gay T., Sholes G.N., and Lieberman Ph. 1968. Some stress effects on the electromyographic measures of consonant articulations. Reports of the 6th International Congress on Acoustics. Ed. Y. Kohasi, Tokyo

Hirano M. and Ohala J. 1967. Use of hooked-wire electrodes for electromyography of the intrinsic laryngeal muscles. Working Papers in Phonetics 7, 35-45. UCLA

Hirano M., Ohala J., and Vennard W. 1969. The function of the laryngeal muscles in regulating fundamental frequency and intensity of phonation. Journal of Speech and Hearing Research 12, 616-628

Hirano M., Vennard W., and Ohala J. 1970. Regulation of register, pitch, and intensity of voice: An electromyographic investigation of intrinsic laryngeal muscles. Folia phoniatrica 22, 1-20

Hirose H., Simada Z., and Fujimura O. 1970. An electromyographic study of the activity of the laryngeal muscles during speech utterances. Annual Bulletin, Research Institute of Logopedics and Phoniatrics, University of Tokyo 4, 9-25

Kamijo Y. 1966. Zusetsukôkû Kaibôgaku, 2 Kingaku (Illustrated Oral Anatomy, 2 Myology). Tokyo: Anatômu-sha

Ohala, J. 1970. Aspects of the Control and Production of Speech. Working Papers in Phonetics 15. UCLA

Ohala J. and Hirose H. 1970. The function of the sternohyoid muscle in speech. Annual Bulletin, Research Institute of Logopedics and Phoniatrics, University of Tokyo 4, 41-44

Sawashima M., Gay T., and Harris K.S. 1969. Laryngeal muscle activity during vocal pitch and intensity changes. Status Report on Speech Research SR-19/20, Haskins Laboratories, 211-220

Simada Z. and Hirose H. 1970. The function of the laryngeal muscles in respects to the word accent distinction. Annual Bulletin, Research Institute of Logopedics and Phoniatrics, University of Tokyo 4, 27-40

Sonesson B. Human vocal folds. To be published

Öhman S.E.G. 1966. Coarticulation in VCV utterances: Spectrographic measurements. JACS 39, 151-168

CONTRASTIVE ACOUSTIC ANALYSIS OF VOWEL PHONEMES, PRONOUNCED BY SOME NORTH GERMAN AND SOUTH SWEDISH HIGH SCHOOL PUPILS

(A summary)

Karin Kitzing

PURPOSE OF THE INVESTIGATION

There are great similarities between the standard Swedish and the German vowel systems. When teaching a foreign language, one must, however, start from the local native dialect of the pupils. One of the characteristics of the Malmö dialect is the pervading diphtongization of the long vowels. The purpose of the present study is to investigate how this dialect influences the pronunciation of the German vowels, spoken by some Malmö pupils.

INFORMANTS

The informants were 107 boys and girls, age 16-21, from a Malmö high school, all representing the Malmö dialect. As a control group the German utterances were recorded by 22 boys of corresponding age at a high school at Lübeck.

MATERIAL

The following 15 vowel phonemes were investigated: /i:, ɪ, e:, ɛ:, ɛ, a:, a, y:, ʏ, ø:, œ, u, ʊ, o:, ɔ./ The test vowels were surrounded by voiceless stops. For the test words, se appendix. The test words were preceded by German: "Das Wort ist..." or Swedish: "Ordet är...". The Swedish pupils read the German sentences first, then the Swedish.

EQUIPMENT

The Swedish material was recorded in an anechoic chamber on a Telefunken M24, and the German in a language laboratory on a Philips RK 65, all at 19.5 cms/sec. The utterances of the male informants were analyzed with a Kay Electric Sona-Graph at the University of Lund. About 1900 spectrograms were made (filter 300 cps, frequency scale 4800 cps).

RESULTS

The median frequencies of the vowel formants for each phoneme and the duration are given in the appendix. Three questions are considered - the frequency of the vowel formants, the duration of the vowels and the centralizing tendency. The last point is related to a quality difference between long tense vowels and the corresponding short lax ones. A perceptual analysis of all informants, including the females, is in progress.

DISCUSSION

1. The frequency of the vowel formants

In spite of the pervading diphthongization of chiefly the long vowels in the Malmö dialect, only few pupils diphthongized the German test vowels. As mentioned above, diphthongization is the main characteristic of the Malmö dialect, and the pupils, well aware of it, avoided diphthongizing the vowels of the target language. Other differences in the pronunciation of the vowels in the Malmö dialect and German are seldom noticed by the pupils or by the teachers of German. The values of the formant frequencies point to the fact that, as a rule, the Malmö informants used their habitual articulatory model when pronouncing the short German vowels. The articulation of the front rounded vowels differed most conspicuously from that of the German informants. Because of the weak labialization of the Malmö dialect, F2 and F3 were higher in the German vowels pronounced by the Malmö informants than in the same vowels articulated by Germans. In the back vowels, the weaker labialization gave a higher F2 in the Swedish pronunciation. Further, the Malmö pupils have too close an /I/ in "kicken", the same pronunciation as they use in the Swedish word "kicka". /ɛ/ is pronounced closer and /ɛ:/ more open than in the native German articulation. All these phenomena might be regarded as interference from the native on the target language.

2. The duration of the vowels

Most of the German vowels are lengthened in the articulation of the Swedish informants. This characteristic alone might bear witness to a foreign pronunciation, even if the frequency of the vowel formants could be representative of German vowels. While the Malmö pupils avoided diphthongizing the tense German vowels, as mentioned above, they appear to lengthen them instead. The Scanian realization of the long Swedish vowels is mainly characterized by two phenomena: diphthongization and lengthening. Of these two features, the Malmö informants only transferred the extra duration to the target language German. As was the case with the formant frequencies, the durations of the Swedish articulations of the German vowels were intermediate between the Swedish pronunciation of the Swedish vowels and the German pronunciation of the German vowels.

3. Centralizing tendency

With centralizing is meant the fact that the tongue hump approaches a neutral mean position in the mouth for the short lax vowels as compared to the more extreme position for the long tense vowels. There is also an acoustic counterpart to the articulatory centralization. The acoustic neutral position is represented by a higher F1 and lower F2 for the short lax front vowels than for the corresponding tense ones. The back rounded lax vowels, however, have higher F2 values than the corresponding tense ones. The present study examines how this acoustic difference is observed by the different informant groups. In almost all cases the frequency difference in F1 between long and short vowels was greater in the German than in the Swedish pronunciations. F1 of the long vowels was generally lower in the German than in the Swedish pronunciation. For the short German vowels, however, the relation is reversed: here F1 is higher in the German than in the Swedish pronunciation. As to the front vowels, the centralizing tendency sometimes is seen in F2 sometimes in F3 in the German articulation. The back vowels /u: – ʊ / and /o: – ɔ / both show greater difference in F2 in the German than in the Swedish pronunciation. The present investigation has proved that the centralizing tencency was not sufficiently observed by the Malmö informants. They seem to have neglected the German quality difference. As the quantity generally is distincitive in Swedish and the quality is regarded as allophonic, there is a natural tendency to carry over this relationship to a foreign language. This case too might be regarded as a kind of interference.

PEDAGOGIC CONCLUSIONS

Some conclusions which could be applied in the teaching of German as a foreign language in southern Sweden could be drawn from the acoustic results presented above:

1. The characteristic strong labialization of the rounded German vowels is important for a good German pronunciation.

2. The pupils should not exaggerate the vowel length.

3. The teacher of German should stress the quality difference between /i: - I, y: - Y, ø: - œ/ and so on and point to the fact that it is not only a question of a quantity difference.

MEDIAN AND RANGE OF VARIATION OF THE FORMANT FREQUENCIES

The numbers in parenthesis indicate the number of informants. Recordings have sometimes been rejected because of reading mistakes, difficulty in measuring spectrograms etc. The number of informants varies therefore from phoneme to phoneme. "sv." = the pronunciation of the Swedes, "ty." = the pronunciation of the Germans. "m." = the vowel in the Swedish word pronounced as a monophthong by the number of informants indicated in parenthesis, "d."= as a diphthong.

	F1	var.bredd	F2	var.bredd	F3	var.bredd
piepen sv. /i:/	325(47)	275–400	2225(47)	1925–2700	2900(46)	2525–3575
pipa	d.(46) m.(4)					
piepen ty.	300(22)	250–350	2200(21)	1950–2500	2863(22)	2650–3150
kicken /I/ sv.	350(49)	300–400	2200(49)	1950–2650	2813(48)	2325–3225
kicka	350(50)	300–400	2238(50)	2000–2725	2875(49)	2575–3225
kicken ty.	350(21)	325–400	2125(21)	1850–2375	2550(21)	2350–2725
Theke* /e:/ sv.	400(49)	300–475	2125(49)	1925–2625	2750(49)	2400–3175
tekopp d.(41)m.	425(9)	400–500	2075(9)	1850–2650	2775(9)	2525–3175
Theke ty.	375(21)	325–450	2150(21)	2025–2425	2700(21)	2550–3000
Täter* /ɛ:/ sv.	550(50)	400–650	1788(50)	1450–2300	2663(50)	2350–3325
täta d.(13)m.	550(37)	450–625	1700(37)	1475–2200	2700(37)	2550–3075
Täter /ɛ:/ ty.	563(3)		1900(3)		2562(3)	
Täter /e:/ ty.	375(19)	300–475	2175(19)	2025–2375	2700(19)	2475–2900
tätscheln* /ɛ/ sv.	500(50)	375–675	1900(50)	1600–2200	2650(50)	2375–3075
tätting d.(35)m.	500(14)	450–550	2000(14)	1625–2450	2675(14)	2550–3125
tätscheln ty.	575(22)	475–625	1863(22)	1625–2050	2575(22)	2350–2850
Pate* /a:/ sv.	700(50)	550–900	1350(50)	1225–1575		
pater /ɑ:/ d.(20)m.	575(26)	450–750	1050(26)	775–1400		
Pate ty.	725(21)	625–850	1200(21)	1050–1300		
packen /a/ sv.	650(50)	500–900	1500(50)	1275–1700		
packa	675(48)	425–875	1500(48)	1250–1700		
packen ty.	750(22)	550–900	1400(22)	1200–1550		
Typus /y:/ sv.	325(48)	275–425	2000(47)	1700–2625	2525(47)	2275–3025
typ	d.(48)m.(2)					
Typus ty.	300(20)	250–350	1738(20)	1500–1950	2200(20)	1900–2375
Stück /ʏ/ sv.	375(50)	300–425	1925(50)	1575–2250	2500(50)	2225–2900
styck	350(49)	300–425	1925(49)	1500–2175	2550(49)	2325–3075
Stück ty.	375(20)	350–475	1500(20)	1375–1625	2200(20)	1925–2500
Spöke* /ø:/ sv.	450(50)	350–575	1575(50)	1375–2125	2488(48)	2225–2975
spöke d.(22)m.	463(28)	400–575	1500(28)	1300–1900	2625(28)	2325–2925
Spöke ty.	400(20)	350–475	1425(20)	1300–1600	2125(19)	1925–2375
Stöpsel /œ/ sv.	500(49)	425–625	1550(49)	1350–1850	2575(49)	2300–2900
stöppla	475(50)	400–575	1525(50)	1300–1850	2638(48)	2400–3050
Stöpsel ty.	550(22)	400–650	1425(22)	1300–1550	2338(22)	1900–2600

* The following table divides the informants into two groups, according to their pronunciation of the Swedish test word. One group diphthongizes the Swedish test vowel, the other group does not.

Appendix 6

Kuhkalb /u:/ sv.	325(44)	300–450	850(44)	675–1125
kok	d.(49)m.(1)			
Kuhkalb ty.	300(22)	250–350	688(22)	575–950
Butter /ʊ/ sv.	350(50)	300–450	950(50)	775–1250
bott	350(48)	300–450	913(46)	600–1250
Butter ty.	425(21)	325–450	875(21)	750–1050
tot /o:/ sv.	425(50)	350–500	850(50)	700–1100
tåt	d.(48) m.(2)			
tot ty.	375(21)	300–450	725(21)	550–800
Pocke /ɔ/ sv.	475(50)	400–600	975(50)	800–1125
pocka	450(50)	325–575	913(50)	700–1125
Pocke ty.	550(20)	475–650	963(20)	850–1075

MEDIAN FORMANT FREQUENCIES AND MEDIAN DURATION OF SOME GERMAN VOWELS DIVIDED ACCORDING TO THE DIPHTHONGIZATION OF THE CORRESPONDING SWEDISH VOWELS

Svenskt testord		Tyskt testord	F1	F2	F3	Duration i msec.
tekopp	monoftong (9)	Theke /e:/	400	2100	2750	158
	diftong (40)		400	2125	2725	158
täta	monoftong (37)	Täter /ɛ:/	525	1800	2650	158
	diftong (13)		550	1750	2700	150
tätting	monoftong (14)	tätscheln /ɛ/	500	1925	2650	90
	diftong (35)		500	1900	2650	98
pater	monoftong (26)	Pate /a:/	700	1350		158
	diftong (20)		700	1325		165
spöke	monoftong (28)	Spöke /ø:/	450	1575	2525	162
	diftong (22)		450	1563	2450	162

/i:/ in "pipa" is pronounced as a monophthong by 4 Swedish informants
/y:/ in "typ" " " " " " " 2 " "
/o:/ in "tåt" " " " " " " 1 " informant.

A division on the basis of these test vowels has not been made because of the very small number of informants pronuncing the Swedish test vowel as a monophthong.

MEDIAN AND RANGE OF VARIATION OF THE VOWEL DURATION IN MSEC

The number in parenthesis indicates the number of informants. Recordings have sometimes been rejected because of reading mistakes, difficulty in measuring spectrograms etc. The number of informants varies therefore from phoneme to phoneme. "sv." = the Swedish pronunciation of the German word, "ty." = the German pronunciation of the German word, "m." = the vowel in the Swedish word pronounced as a monophthong, "d." = as a diphthong.

piepen	/i:/	sv.	150	msec.	(47)	var.	bredd.	105–218	msec.
pipa		d.	165	"	(46)	"	"	113–225	"
piepen		ty.	83	"	(22)	"	"	53–128	"
kicken	/I/	sv.	90	"	(50)	"	"	45–135	"
kicka			105	"	(50)	"	"	53–150	"
kicken		ty.	60	"	(21)	"	"	30– 90	"
Theke	/e:/	sv.	158	"	(49)	"	"	105–210	"
tekopp		m.	150	"	(9)	"	"	105–195	"
tekopp		d.	150	"	(41)	"	"	120–225	"
Theke		ty.	120	"	(21)	"	"	98–158	"
Täter	/ɛ:/	sv.	158	"	(50)	"	"	90–233	"
täta		m.	158	"	(37)	"	"	188–195	"
täta		d.	165	"	(13)	"	"	135–225	"
Täter	/e:/	ty.	113	"	(19)	"	"	83–150	"
tätscheln	/ɛ/	sv.	94	"	(50)	"	"	68–180	"
tätting		m.	98	"	(14)	"	"	68–135	"
tätting		d.	113	"	(36)	"	"	68–158	"
tätscheln		ty.	68	"	(22)	"	"	53– 90	"
Pate	/a:/	sv.	158	"	(50)	"	"	105–203	"
pater	/ɑ:/	m.	158	"	(26)	"	"	113–188	"
pater		d.	173	"	(19)	"	"	120–218	"
Pate		ty.	150	"	(22)	"	"	128–188	"
packen	/a/	sv.	105	"	(50)	"	"	68–173	"
packa			120	"	(50)	"	"	68–195	"
packen		ty.	68	"	(22)	"	"	53–105	"
Typus	/y:/	sv.	135	"	(49)	"	"	83–210	"
typ		d.	188	"	(48)	"	"	120–248	"
Typus		ty.	60	"	(21)	"	"	45–98	"
Stück	/Y/	sv.	120	"	(50)	"	"	75–165	"
styck			143	"	(49)	"	"	90–203	"
Stück		ty.	90	"	(20)	"	"	68–143	"
Spöke	/ø:/	sv.	162	"	(50)	"	"	113–210	"
spöke		m.	173	"	(28)	"	"	128–210	"
spöke		d.	188	"	(22)	"	"	135–225	"
Spöke		ty.	132	"	(20)	"	"	105–158	"
Stöpsel	/œ/	sv.	98	"	(49)	"	"	60–128	"
stöppla			120	"	(50)	"	"	75–150	"
Stöpsel		ty.	71	"	(22)	"	"	53– 90	"

Appendix 8

Kuhkalb	/u:/	sv.	120 msec.	(49) var. bredd.			75–225 msec.
kok		d.	188 "	(49)	"	"	105–233 "
Kuhkalb		ty.	79 "	(22)	"	"	53–120 "
Butter	/ʊ/	sv.	113 "	(50)	"	"	53–158 "
bott			150 "	(48)	"	"	105–203 "
Butter		ty.	75 "	(21)	"	"	45–105 "
tot	/o:/	sv.	188 "	(50)	"	"	120–278 "
tåt		d.	195 "	(48)	"	"	128–255 "
tot		ty.	173 "	(21)	"	"	135–218 "
Pocke	/ɔ/	sv.	94 "	(50)	"	"	30–135 "
pocka			98 "	(50)	"	"	53–135 "
Pocke		ty.	64 "	(20)	"	"	45– 90 "

A SPECTROGRAPHIC STUDY OF ALLOPHONIC VARIATION AND VOWEL REDUCTION IN WEST GREENLANDIC ESKIMO*

Sidney Wood

SUMMARY

West Greenlandic Eskimo vowel spectra have been investigated in carefully pronounced words and in continuous speech, and spectral differences observed between the two situations. Spectra have been compared in two consonantal environments - pharyngeal (uvular) and non-pharyngeal. In order to reconstruct vowel articulations, reference has been made to the three-parameter model of vowel production. It is inferred from this that stressed vowels in non-pharyngeal environments would require pharyngeal to velar constrictions for /a/, velar for /u/ and palatal for /i/. For the allophones in pharyngeal environments, the necessary constrictions would be low-pharyngeal for /a/ and uvular for /i/ and /u/. Fully reduced vowels in non-pharyngeal environments would have uvular to velar articulations with fairly narrow degrees of constriction, and in pharyngeal environments low-pharyngeal to uvular articulations with a narrower degree of constriction. In both cases, the mouth-opening would have been less than moderate. The regression of vowel spectra from target to reduced is apparently mainly associated with a narrowing of the range of mouth-openings, while there can be some centralization of constriction locations towards the velar region provided the degree of constriction remains small. An approximation to the uniform tube does not seem to have been a likely configuration for this informant's weak vowels.

* I am indebted to Jörgen Rischel and Carl-Christian Olsen of Copenhagen University for practical assistance and for valuable discussions about WG phonology.

1. INTRODUCTION

1.1 West Greenlandic Eskimo

1.1.1 W.G. Eskimo has three vowel phonemes, /a, i, u/, but a fourth abstract morpho-phoneme /ə/ can be posited to explain certain morphological alternations. This abstract unit is expressed either as [i] or [a], merging with the reflexes of /i/ and /a/. This investigation is based on the three-phoneme solution, but it will not have biased the results to ignore /ə/ since its reflexes have no separate identity in speech. W.G. has four contrasting articulations for stops, from front to back /p, t, k, q/. There has been some dispute about the contrasting articulations of the back pair – velar/uvular or palatal/velar. There is a back fricative /ʁ/ corresponding to /q/. The informant's pronunciation of /q, ʁ / sounded very retracted, probably below the velum, and will be referred to as "pharyngeal". The question of whether it was at or below the uvula will be left open. A short description of Eskimo is given by Hill (1958: Appendix A).

1.1.2 When vowels precede [q] or [ʁ] , the body of the tongue is retracted in anticipation of the consonant articulation. For /a/ and /u/, this results in cardinal 5 [ɑ] and cardinal 6 or 7 [ɔ] – [o] respectively, while palatal /i/ is said to be uvularized and sounds like a retracted cardinal 3, [ɛ̈] , or the plain back cardinal 14 or 15, [ʌ] – [ɤ]. The exact description of this allophone has been a matter of controversy. It has been transcribed by Thalbitzer as [ɜ] and by Uldall and Lawrenson as [ᵊ] (Lawrenson 1934).

1.1.3 Investigation of the spectra of these vowels may to some extent clarify the problem of the nature of the vowel allophones before [q] and [ʁ] .

1.2 Transcription

The vowel allophones will be denoted as follows:

position	/a/	/i/	/u/
In pharyngeal environments (before [q] or [ʁ])	ɑ	ɨ	o
In non-pharynageal environments (before consonants with any other place of articulation than [q] or [ʁ])	a	i	u

1.3 The Investigation

The investigation is devoted to the following topics:

(a) § 4.1 deals with allophonic variation. The spectral and inferred articulatory differences between vowels in the pharyngeal and non-pharyngeal environments are considered separately for three different situations: (i) stressed vowels in carefully pronounced words at § 4.1.1, (ii) stressed vowels in contiuous speech at § 4.1.2, (iii) weak vowels in continuous speech at § 4.1.3. Articulations are inferred from the three-parameter model of vowel production (cf. § 2.5).

(b) § 4.2 deals with vowel reduction. The progressive centralization of spectra from stressed to weak vowels is considered separately for the two major sets of allophones. Inferred articulations for the spectra of fully reduced vowels are given in § 4.2.2. The degree of correlation between the first two formants of different renderings of a phoneme is examined in § 4.2.3. In particular, the articulatory centralization associated with the spectral centralization along the regression of the formant frequencies for each phoneme is discussed in § 4.2.3 (c, d).

(c) § 4.3 brings together the inferred articulations, and a rule is proposed to generate [ɑ], [ɨ] and [o] from /a/, /i/ and /u/ respectively in pharyngeal environments. The place of articulation of the pharyngeal consonants, hitherto an open question in this paper is also discussed.

2. PROCEDURE

2.1 Informant and Recordings

The informant is an adult male native speaker of the Holsteinsborg dialect. He has recorded a set of carefully pronounced isolated words with stressed vowels in pharyngeal and non-pharyngeal environments, and a passage of continuous speech (a few pages from the novel "Singnagtugaq" by Mathias Storch). The recordings were made in the phonetics laboratory of Copenhagen University. The analysis was carried out at Lund with a Voiceprint spectrum analyser.

2.2 Speech material

2.2.1 Single word utterances

The carefully pronounced words contained 30 stressed vowels, as follows:

	/a/	/i/	/u/
non-pharyngeal environments	7	6	5
pharyngeal environments	4	4	3

2.2.2 Continuous speech

(a) Analysis of the continuous speech was broken off after 262 phonetic syllables. This gave 250 syllables containing vowel spectra, while the remaining 12 were rejected because their nuclei were carried by consonants, for example:

...pisinialeramik...

[pəʒ-n̥-nje-ˈlɔʁ-əm-ək]

"as they wanted themselves to do business"

cf. spectrogram at Fig. 1 a.

The informant's reading was very informal and the speaking rate rather fast - an average of 375 syllables per minute. The passage of continuous speech is consequently characterized by considerable contraction of syllables. For example:

(i) ...kînáinitdlo nalunángileq ingmingnut...

[kɪ-ˈnaj-nəl̥-l̥o-ə-ˌlaŋ-ə-ˈlɪm-mɪn-ət]

"and from their faces/was obvious/to themselves"

(ii) ...qimáinardlugit uvagut ingmíkut...

[qə-ˌmajn-nal̥-ˌl̥ju-wɪm-məkt]

"simply leaving them/we/ourselves"

cf. spectrograms at Fig 1 (b, c)

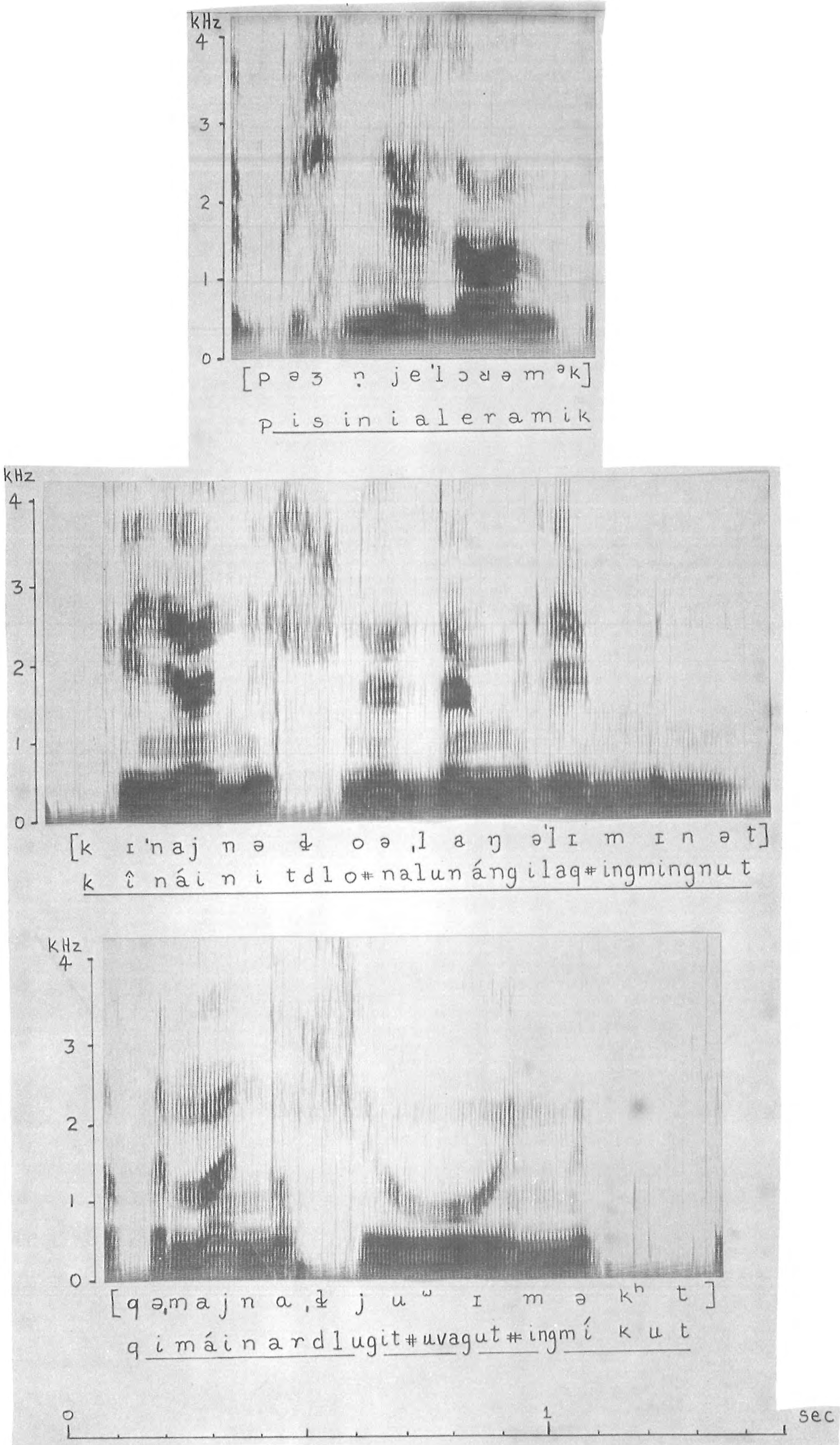

Figure 1. Examples of syllable contraction in continuous speech

(b) the 250 syllables contained the following vowels:

	/a/	/i/	/u/
Strong, non-pharyngeal	25	7	18
Weak, non-pharyngeal	48	52	38
Strong, pharyngeal	15	8	2
Weak, pharyngeal	10	13	14

(c) Many of these vowels were severely distorted by the contraction of syllables. The following smaller set of vowels is taken from syllables that were without any doubt undistorted:

	/a/	/i/	/u/
Strong non-pharyngeal	22	7	16
Weak non-pharyngeal	14	17	12
Strong pharyngeal	15	8	2
Weak pharyngeal	7	11	9

These are the vowels used for Figs. 2 (b,c,e,f).

2.3 Formant frequencies

The spectra of all vowels were sampled once only at a point in time where they were presumed most closely to approach a supposed target spectrum (denoted F_{1o}, F_{2o} ... F_{no}), cf. Fig 2.

2.4 Stress judgements

Levels of phonetic stress have been judged subjectively against a three-grade scale. The strong and medium grades are pooled in the results. The numbers of stressed and weak vowels have already been given above, § 2.2.

The stress levels refer to sentence stress in the informant's speech, and not to any abstract underlying lexical stress categories.

2.5 Inferred articulations

In an attempt to reconstruct articulations for the observed spectra, recourse is made to the three-parameter model of vowel production (Chiba and Kajiyama 1941, Fant 1960 and 1968, Stevens and House 1955 and 1961). According to this model, the resonator configuration can be specified by (i) the cross-section area at the narrowest constriction, A_o cm^2, (ii) the distance of that point from the source, d_o cms, and (iii) the opening area/length ratio, A/l cms. Particular reference is made to the data given by Stevens and House (1955: Fig. 5) for contours of constant formant frequency for different magnitudes of the parameters in an acoustic tube simulation of the typical adult male vocal tract. This version of the model has the tube radius at the constriction, r_o cm., rather than the cross-section area πr_o^2. Vocal tract cross-section areas corresponding to tube constriction radii are:

Constriction radius in the tube	r_o	0.3 cm	0.6 cm	0.8 cm	1.0 cm
Corresponding cross-section area in the vocal tract	A_o	0.3 cm^2	1.1 cm^2	2.0 cm^2	3.1 cm^2

The following points must be borne in mind when inferences are made from this model:

(a) The model assumes a constant tube length. In human speech, a palatal gesture of the tongue lifts the larynx via the hyoid bone. Consequently, the vocal tract is shorter for palatal vowels, especially for [i]. This might influence the articulations inferred for palatal vowels.

(b) An unknown quantity in this type of reconstruction is the magnitude of the cross-section area A_o. Several alternative solutions must therefore be compared, for selected values of this parameter. It is often possible to set a maximum permissible limit beyond which it is impossible to generate given spectra. The effect of increasing A_o is gradually to neutralize the parameter d_o and to reduce the total possible spectral variation. Fant 1960 (Table 2.33-1) gives the following constriction cross-section areas for a set of Russian vowels:

0.65 cm^2 [ɑ, i] 1.0 cm^2 [o, u, ɨ] 2.0 cm^2 [e]

This same range of cross-section areas will be used in the discussion, except that the lower limit will be set at Steven and House's smallest degree r_o = 0.3 cm (A_o = 0.3 cm^2).

For palatal constrictions, and to some extent velar constrictions, we can expect the degree of constriction to be a function of jaw-opening (cf. the data quoted above for the Russian vowels where A_o was 0.65 cm^2 for /i/ and 2.0 cm^2 for /e/), but for pharyngeal constrictions, the degree of constriction can be small even with a large jaw-opening (A_o for the Russian /ɑ/ was 0.65 cm^2).

(c) The model does not concern itself with the articulatory degrees of freedom of the human vocal tract, and it is therefore capable of making combinations of parameter values that would hardly occur in natural speech. The articulatory degrees of freedom are discussed by Lindblom and Sundberg (1969: § 2). The same authors have also investigated the acoustic consequences of movements of the lip, tongue, jaw and larynx (1971). In the present investigation, [u] spectra are found which could be generated with a constriction low in the pharynx, or [i] spectra with an extremely large mouth-opening. Our knowledge of general phonetics would say that such articulations are very unlikely for these vowels and would limit the choice of acceptable inferences.

(d) The data given by Stevens and House is for a resonator with typical adult male dimensions. But it is not certain that the informant is representative of the typical case. The pitches of his second formant were almost normally distributed about a mean of 1200 mels (1325 Hz) when sampled at 0.025 second intervals through the continuous speech. The mean of this distribution differs individually between speakers and is possibly related to vocal tract dimensions. If so, this informant's mean F2 pitch (which is on the low side) might indicate that he has a larger-than-average vocal tract. There is then a risk that the model will underestimate the distance from source to contriction for this informant, especially when this distance is large - i.e. for palatal constrictions.

3. RESULTS

3.1 Fig. 2 (a-f) shows the F_{1o}/F_{2o} areas of stressed vowels in the single word utterances, and stressed and weak vowels in continuous speech, for post-vocalic non-pharyngeal and pharyngeal environments. The areas represent the full variation observed for each phoneme - the only factors taken into account have been speaking situation, stress and the pharyngeal environment, for which spectra have been plotted separately.

3.2 Fig. 3 shows the frequencies F_{1o}, F_{2o} and F_{3o} of the stressed vowels in single word utterances, ranked by ascending F_{1o}, for the non-pharyngeal and pharyngeal environments. There are the same 30 vowels as for Fig 2 (a, d). Fig. 4 shows the frequencies F_{1o}, F_{2o} and F_{3o} of stressed vowels in continuous speech, ranked by ascending F_{1o}, for non-pharyngeal and pharyngeal environments. Fewer vowels were available here than for Fig. 2 (b, e) because F3 was sometimes weak and did not always register on spectrograms.

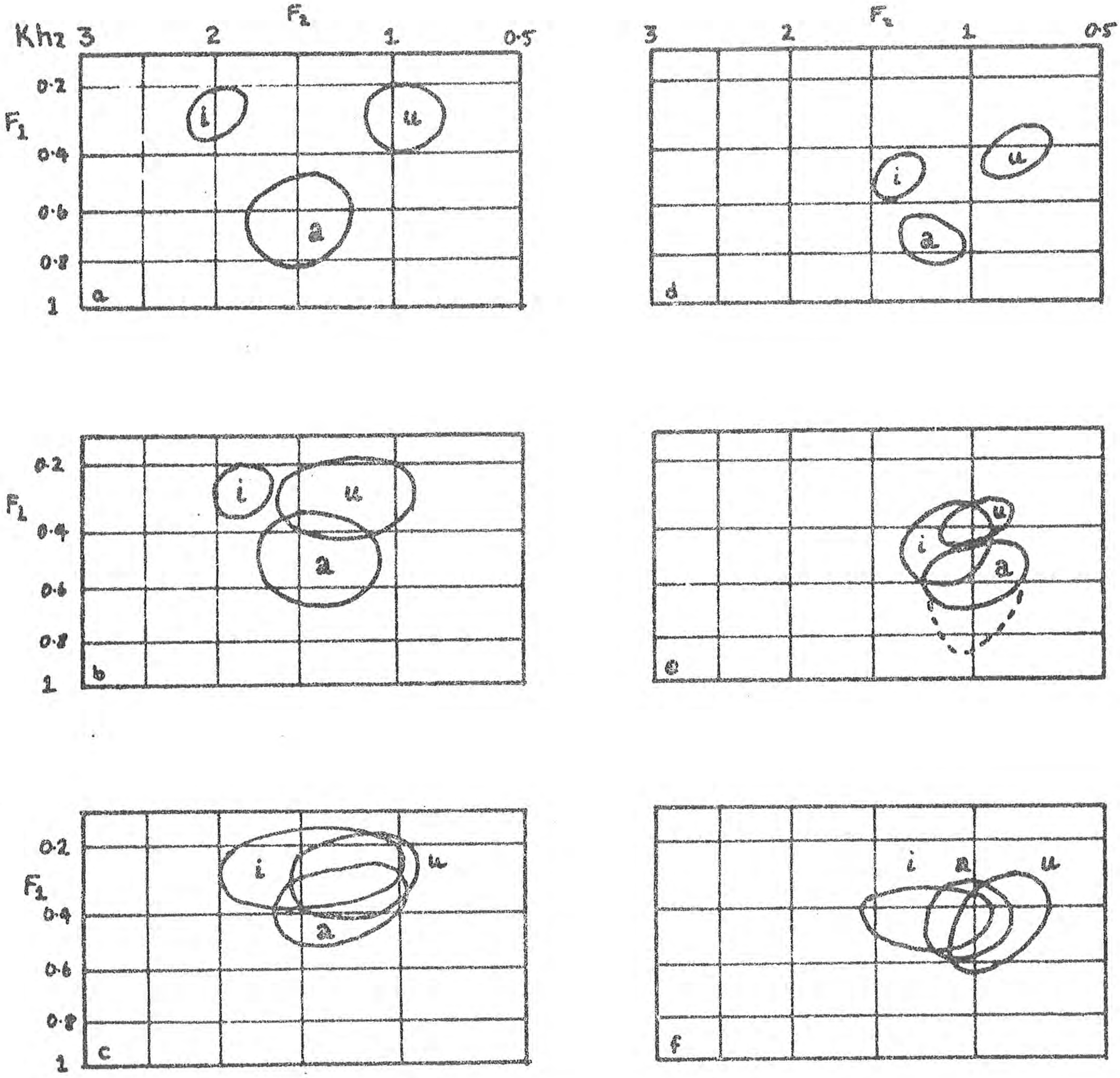

Figure 2. F_{1o}/F_{2o} areas of stressed vowels in carefully pronounced single word utterances (a, d), stressed vowels in continuous speech (b, e) and weak vowels in continuous speech (c, f), in non-pharyngeal(a, b, c) and pharyngeal (d, e, f) environments.

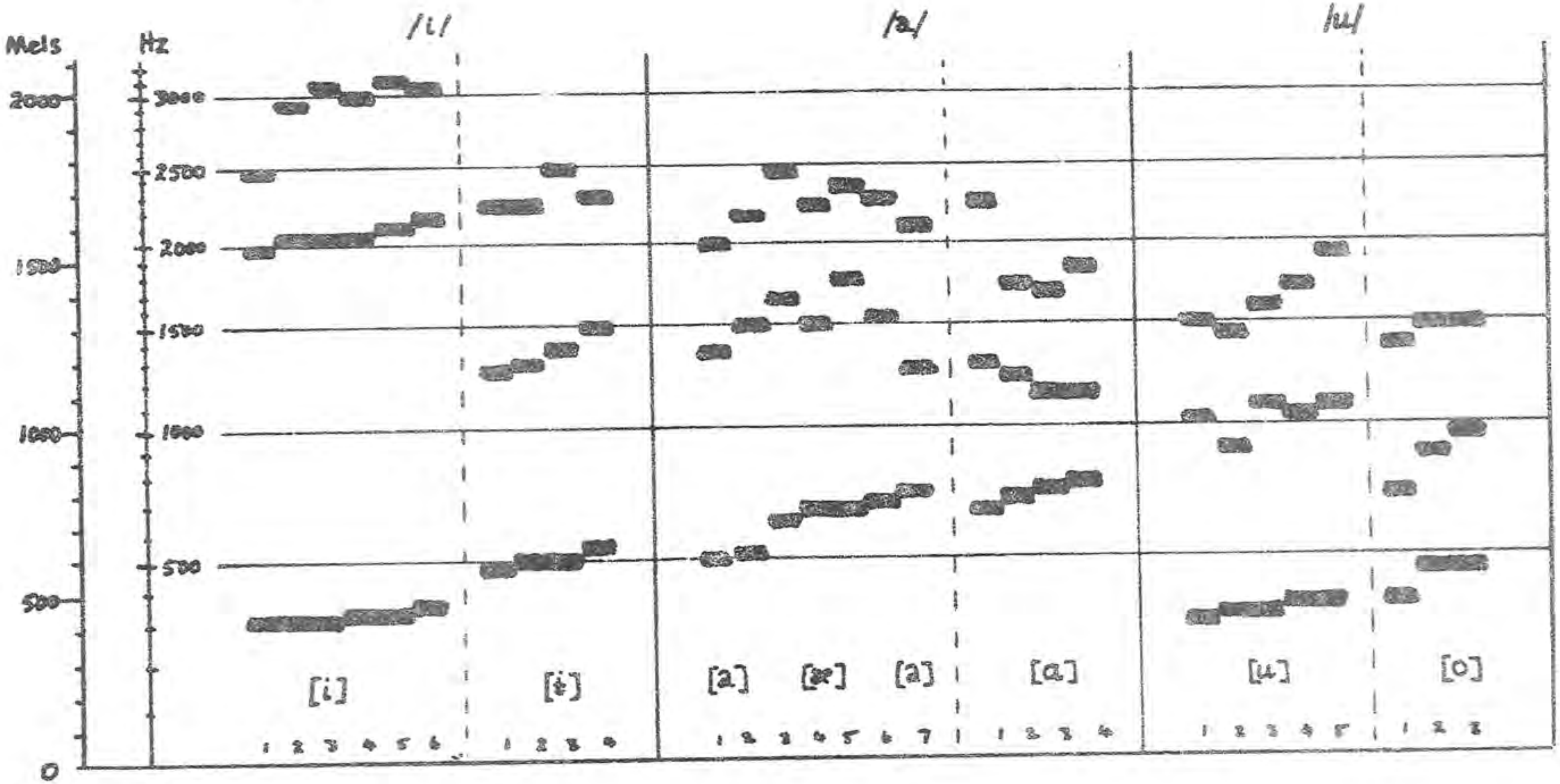

Figure 3. F_{1o}, F_{2o} and F_{3o} of stressed vowels in carefully pronounced single word utterances.

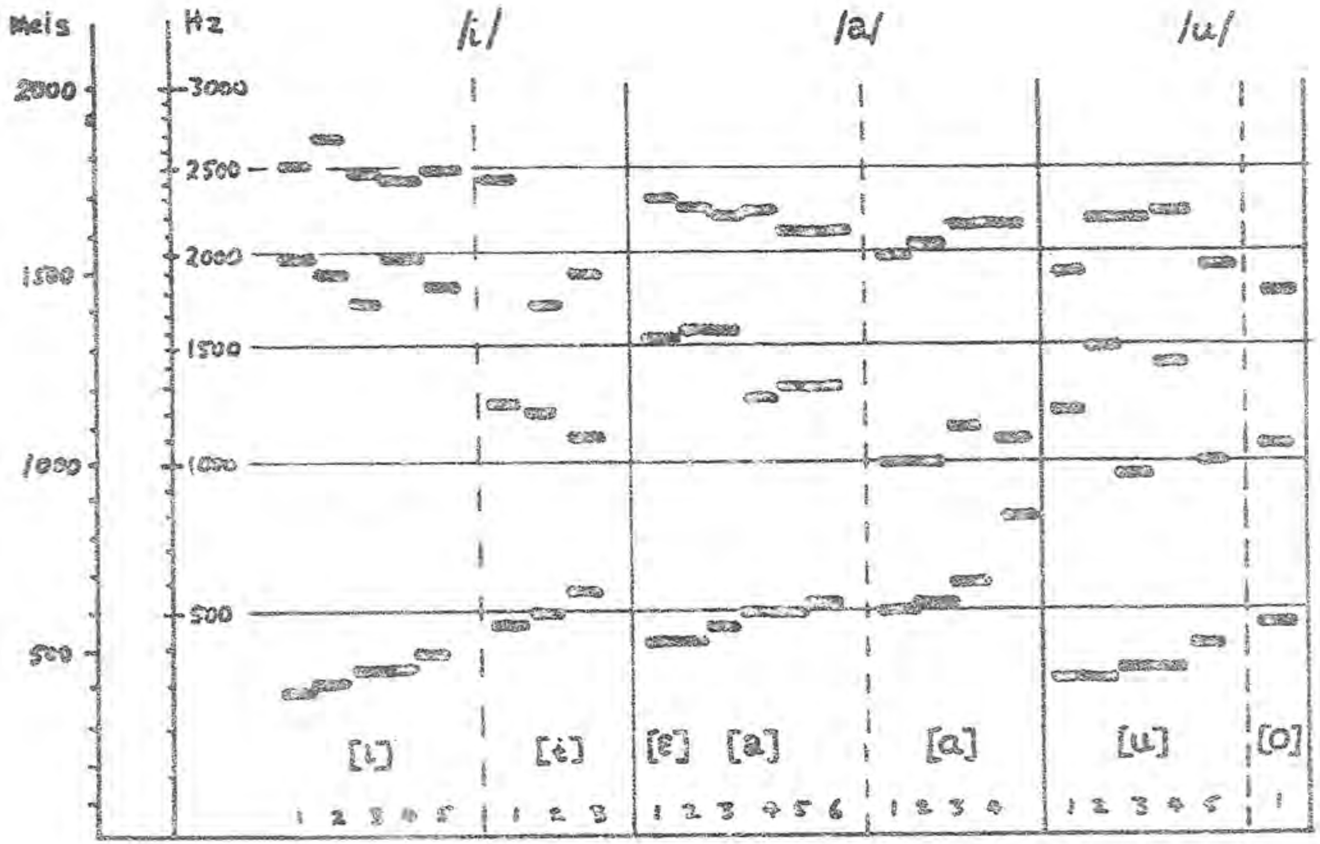

Figure 4. F_{1o}, F_{2o} and F_{3o} of stressed vowels in continuous speech.

4. DISCUSSION AND CONCLUSIONS

4.1 Allophonic variation

4.1.1 Single word utterances

(a) Fig. 2 (a,d) shows that the three phonemes were spectrally well separated in both environments. In non-pharyngeal environments, these spectra were peripheral to the total area used. In pharyngeal environments F_{2o} was lower for all phonemes, and F_{1o} higher for /i/ and /u/.

(b) Fig. 3 shows the frequencies of the first three formants of the distinctive spectra of stressed vowels in the single word utterances, ranked by ascending F_{1o}. For allophones of /i/ and /u/, there was a clear drop of F_{2o} and F_{3o} between the non-pharyngeal and pharyngeal allophones. For /a/, however, there appears to be a continuum with no abrupt spectral gap between the allophones. It is interesting to note that WG orthography distinguishes between the allophones of /i/ and /u/, but not between those of /a/ (i - e, u - o, a).

(c) What are the articulations that can be inferred for these peripheral spectra? Fig. 5 a gives the place of constriction and relative mouth-opening for the non-pharyngeal environments, for three degrees of constriction. A small constriction ($A_o = 0.3\ cm^2$) would generate all the spectra. With wider constrictions, some of the spectra are cut off. For the intermediate case, $A_o = 1.1\ cm^2$, many of the possible /u/ spectra would require constrictions low in the pharynx ($4 < d_o < 7$ cms.) which we might intuitively feel to be unlikely. Similarly, some of the /i/ spectra would require exceptionally large compensatory mouth-openings. In the largest case quoted, $A_o = 2.0\ cm^2$, none of the /i/ spectra would be possible and much of /u/ is still improbably low in the pharynx. It seems almost certain that these two

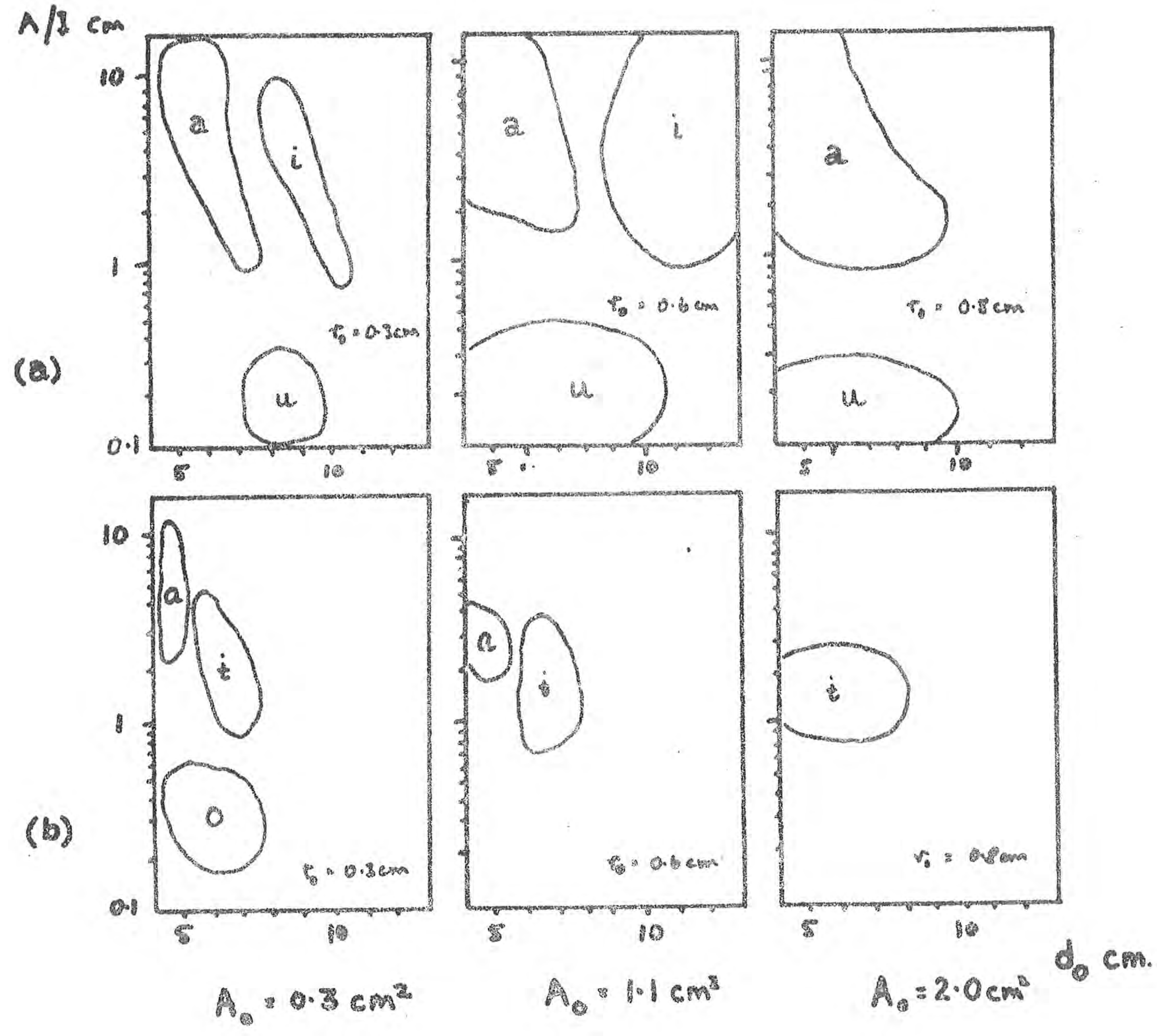

Figure 5. Contours of articulation for the informant's stressed vowel spectra in carefully pronounced single word utterances: (a) non-pharyngeal environments, and (b) pharyngeal environments, for three different degrees of constriction. Based on Stevens and House's contours of constant formant frequency (1955, Fig. 5).

phonemes required small constrictions, /u/ possibly more so than /i/. For /a/, however, while the small constriction is essential for some spectra, the effect of wider constrictions is to extend the range of constriction locations up to the velum ($4 < d_o < 10$ cms.) and to increase mouth-openings to the maximum. These are quite conceivable articulations for for the vowel qualities [a] and [æ] (cf. Fig. 3). The place of articulation and mouth-opening can be summarized as follows:

	/a/	/u/	/i/
distance above the glottis	5–8 cms.	7–10 cms.	8–11 cms.
region	pharyngeal uvular	uvular velar	(velar) palatal
mouth-opening	moderate very large	small	moderate large

The constrictions for /i/ may seem unusually far back. Possible explanations for this are (i) that the point of reference, the glottis, is raised for [i] [cf. § 2.5 (a)], (ii) the informant may have a longer-than-average vocal tract [cf. § 2.5 (d)], and (iii) a slightly wider degree of constriction for /i/ than the minimum, intermediate between 0.3 cm^2 and 1.1 cm^2, would give a location higher up the palate (Fig. 5a shows that increasing A_o has the effect of shifting the [i] constriction nearer the opening).

(d) Fig. 5 b shows the corresponding articulations for the pharyngeal environments. The degree of constriction would appear to be even more critical here. When $A_o > 0.3$ cm^2, the /u/ spectra become impossible, and the /a/ spectra cannot be generated when $A_o > 1.1$ cm^2. The /i/ spectra have very similar articulations for $A_o = 0.3$ cm^2 and $A_o = 1.1$ cm^2, but at $A_o = 2.0$ cm^2 the constriction would go deep into the pharynx. Otherwise, generally smaller degrees of constriction seem to be necessary in this environment than in the non-pharyngeal environment. The place of articulation and mouth-opening can be summarized as follows:

	/a/	/u/	/i/
distance above the glottis	4-5 cms.	5-8 cms.	6-8 cms.
region	low-pharyngeal	pharyngeal uvular	uvular
mouth-opening	moderate large	small	moderate

This means that the place of constriction has been displaced 2 - 3 cms. towards the glottis, in the pharyngeal environments, relative to the non-pharyngeal environments. The displacement along the vocal tract is probably a little greater since the larynx is itself falling back.

4.1.2 Stressed vowels in continuous speech

(a) Fig. 2b shows that the spectra used in stressed non-pharyngeal environments were nearer the centre of the spectral space than those in the single word utterances. The spectra fill out the whole central area, but there is only insignificant overlapping. In the pharyngeal environments (Fig. 2a) there is considerable overlapping of phonemes. The broken lines on Fig. 1e enclose one isolated rendering of /a/ - the remainder were concentrated to the area within the unbroken line.

(b) Fig. 4 gives the frequencies F_{1o}, F_{2o} and F_{3o} of stressed vowels in continuous speech, ranked by ascending F_{1o}. As in Fig. 3, F_{1o} is higher in the pharyngeal environments than in the non-pharyngeal environments, for any one phoneme. F_{2o} was still lower for [ɨ] than for [i], but F_{3o} was not consistently lower (as it was in Fig. 3). There was only a slight difference between the allophones of /u/. But the [o] allophone of /u/ was hardly spectrally distinct from the [ɨ] allophone of /i/ - in the cases where overlapping of F_{1o} and F_{2o} has been observed, there was hardly any contrastive power in F_{3o}.

(c) Fig. 6 a gives the articulatory contours for the stressed vowels in non-pharyngeal environments, for the same three degrees of constriction as in Fig. 5. These spectra can all be generated with constrictions up to $A_o = 1.1$ cm^2, which is a more generous limit than was the case for the single word utterances, § 4.1.1 (c). Much of the spectral area of /i/ is cut off at $A_o = 2.0$ cm^2. As was the case in the single word utterances, the /u/ spectra would seem to require improbably low pharyngeal constrictions when $A_o > 0.3$ cm^2. On the other hand, the gestures for /i/ and /a/ spectra at wider constrictions are plausible - higher up the palate for /i/ and towards the velum and back of the palate for /a/. The mouth-openings are much the same at each degree of constriction. For two degrees of constriction, the articulations can be summarized as follows:

(i) $A_o = 0.3$ cm^2	/a/	/u/	/i/
distance above the glottis	5-8 cms.	7-10 cms.	8-11 cms.
region	pharyngeal uvular	uvular velar	(velar) palatal
mouth-opening	moderate large	small	moderate

(ii) $A_o = 1.1$ cm^2	/a/	/u/	/i/
distance above the glottis	5-10 cms.		10-13 cms.
region	pharyngeal velar		palatal

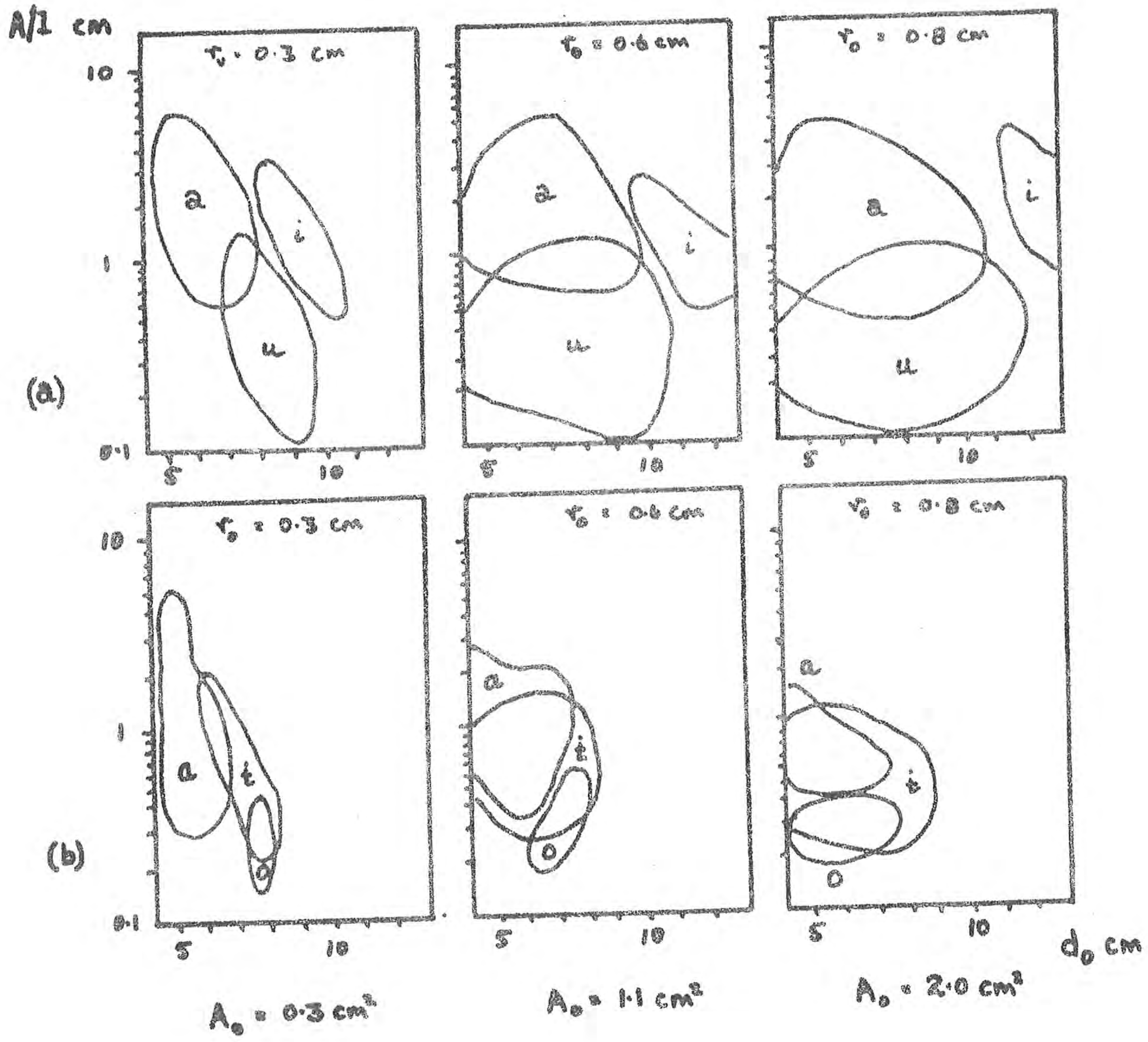

Figure 6. Contours of articulation for the informant's stressed vowel spectra in continuous speech: (a) non-pharyngeal environments, and (b) pharyngeal environments, for three different degrees of constriction.

These are much the same articulations as were found above for the single word utterances. The main differences are that mouth-openings were more moderate in the continuous speech (smaller for /i/ and /a/, often larger for /u/), and that slightly larger degrees of constriction can apparently be tolerated for the continuous speech spectra.

(d) Fig. 6b shows the corresponding articulations for the pharyngeal environments. The overlapping of phoneme areas is equally evident here. The spectra that do not overlap, and which consequently still contrast require the smallest degree of constriction. At larger degrees of constriction, virtually only the overlapping spectra can be generated. For contrasting spectra, the place of articulation and mouth-opening can be summarized as follows:

	/a/	/u/	/i/
distance above the glottis	4–6 cms.	7–8 cms.	6–8 cmc.
region	low-pharyngeal	uvular	uvular
mouth-opening	moderately small	small	moderately small
	moderately large		

4.1.3 <u>Weak vowels in continuous speech</u>

Figs. 1c and 1f show that there was considerable overlapping of the three phonemes in weak syllables in both environments. Comparison of the two charts shows that the spectra of weak vowels in pharyngeal environments are still displaced to a separate area, with generally higher F_{1o} and lower F_{2o} than those in non-pharyngeal environments. Spectral contrasts are largely neutralized, but it would be an over-simplification to say that each phoneme has one weak allophone [ə] in non-pharyngeal environments and another [ɔ] in pharyngeal environments. Fig. 2 (c,f) reveals some semblance of a system, even though it is largely confused by overlapping.

4.2 Vowel reduction

4.2.1 Spectral contrasts

The F_{1o}/F_{2o} areas of the stressed and weak vowels in continuous speech have already been seen at Fig. 2 (b, c, e, f). The focal point for vowel reduction in non-pharyngeal environments was not in the centre (towards say a neutral spectrum of 500, 1500, 2500 ... etc. Hz) but higher up towards a spectrum of 350, 1250 Hz. Many of the weak vowels, whatever the underlying phoneme, did in fact sound like [ʊ] or [o]. The areas for weak /i/ and /a/ extend from this focus of vowel reduction towards the respective stressed areas. The weak pharyngeal allophones had lower F_{1o} than the corresponding strong renderings.

4.2.2 Inferred articulations

(a) Fig 7a gives the contours of articulation for the area of complete overlapping of the weak vowels in non-pharyngeal environments (Fig. 2c). This is the area where vowels are completely reduced and distinctions abandoned. Constrictions would be located 7-10 cms. above the glottis, or roughly from the uvula to the velum with the smallest degree of constriction, $A_o = 0.3$ cm^2. With wider degrees of constriction, a larger range of constriction locations becomes necessary to generate all the spectra in the defined area. When $A_o = 2.0$ cm^2 these spectra are theoretically just possible if constrictions are located from 5 to 12 cms. above the glottis – i.e. utilizing virtually the whole vocal tract. If constrictions are

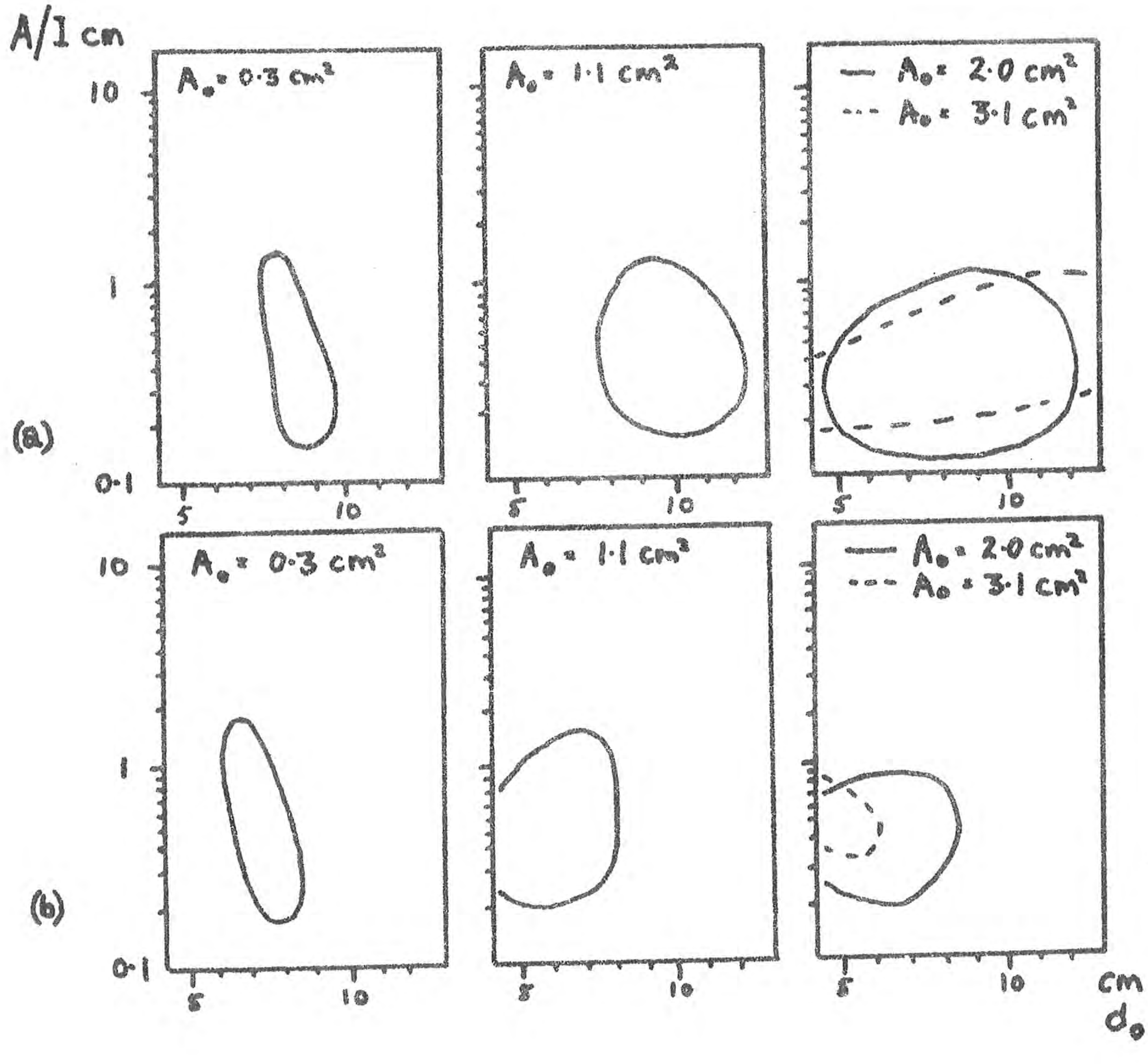

Figure 7. Contours of articulation for the informant's weak vowel spectra that were common to all three phonemes [cf. Fig. 2 (c, f)]: (a) non-pharyngeal environments, and (b) pharyngeal environments, for four degrees of constriction.

even wider, some of the spectra are impossible to generate. There are thus two limiting situations - either (i) constrictions located in all parts of the tract but with cross-section areas up to 1.5 - 2.0 cm^2, or (ii) centralized constrictions located in the velar region with small cross-section areas.

The necessary mouth-openings vary little between the degrees of constriction quoted ($0.1 < A/l < 1.5$ cm), but the upper limit is lower than for stressed vowels. The mouth-opening tended to be smaller for weak vowels than for stressed vowels.

(b) For pharyngeal environments, Fig. 7b shows that the constriction, when very small, is located about 6 - 8 cms. above the glottis, or near the uvula (somewhat lower than for the non-pharyngeal environments). For wider constrictions, the extra range of d_o is only downwards into the pharynx - there is no extension upwards into the palate. Further, some of the spectra are cut off when $A_o > 1.0$ cm^2.

(c) This comparison not only indicates that the tongue constrictions of fully reduced vowels were deeper in the pharynx before the pharyngeal consonants, but also that there was less freedom for constriction size - in pharyngeal environments A_o must be smaller than in non-pharyngeal environments. It is also clear that many of the fully reduced weak vowel spectra could not be generated in a resonator approximating the uniform tube (in Stevens and House's version of the model, the uniform tube has $r_o = 1.2$ cm., corresponding to $A_o = 4.5$ cm^2).

4.2.3 Correlation between F_{1o} and F_{2o}

(a) In a sample of English continuous speech (description forthcoming) it was found that there was a fairly good correlation between F_{1o} and F_{2o}

for different renderings of a phoneme in a homogeneous environment. The regressions for different phonemes were focussed on the speaker's schwa spectra and appeared to indicate the paths of reduction for phonemes in a given environment. It is unfortunately difficult to find a sufficient number of renderings in a purely homogeneous environment in the present WG Eskimo sample of continuous speech, with only 250 syllables to choose from. However it did prove possible to extract a small set of syllables - weak and strong - that excluded labial, palatal, pharyngeal and lateral environments - i.e. a mixture of dental and velar environments remained. A further condition imposed on the selection was that these syllables should faithfully reproduce the underlying forms (phonetic syllables such as those given as examples of syllable contraction in § 2.2 (and Fig. 1) can contain features from a number of underlying syllables and are therefore open to several sources of contamination).

(b) Fig. 8a shows the number of renderings, the product-moment correlation coefficient between F_{1o} and F_{2o}, and the regression of F_{1o} on F_{2o} for each of the phonemes.* These regressions can be compared with the areas of full variation in all environments at Fig. 2 (a-c).

---

* The correlations and regressions given at Fig. 8a must be accepted with some caution because they are based on such small samples of vowels. A t-test gives the following significance levels (p) for these correlation coefficients (r) and sample sizes:

	/i/	/a/	/u/
vowels	10	7	6
r	-0.4	+0.6	+0.3
p	15 %	8 %	30 %

Nevertheless, they do follow the tendency revealed in the much larger sample of English continuous speech referred to above.

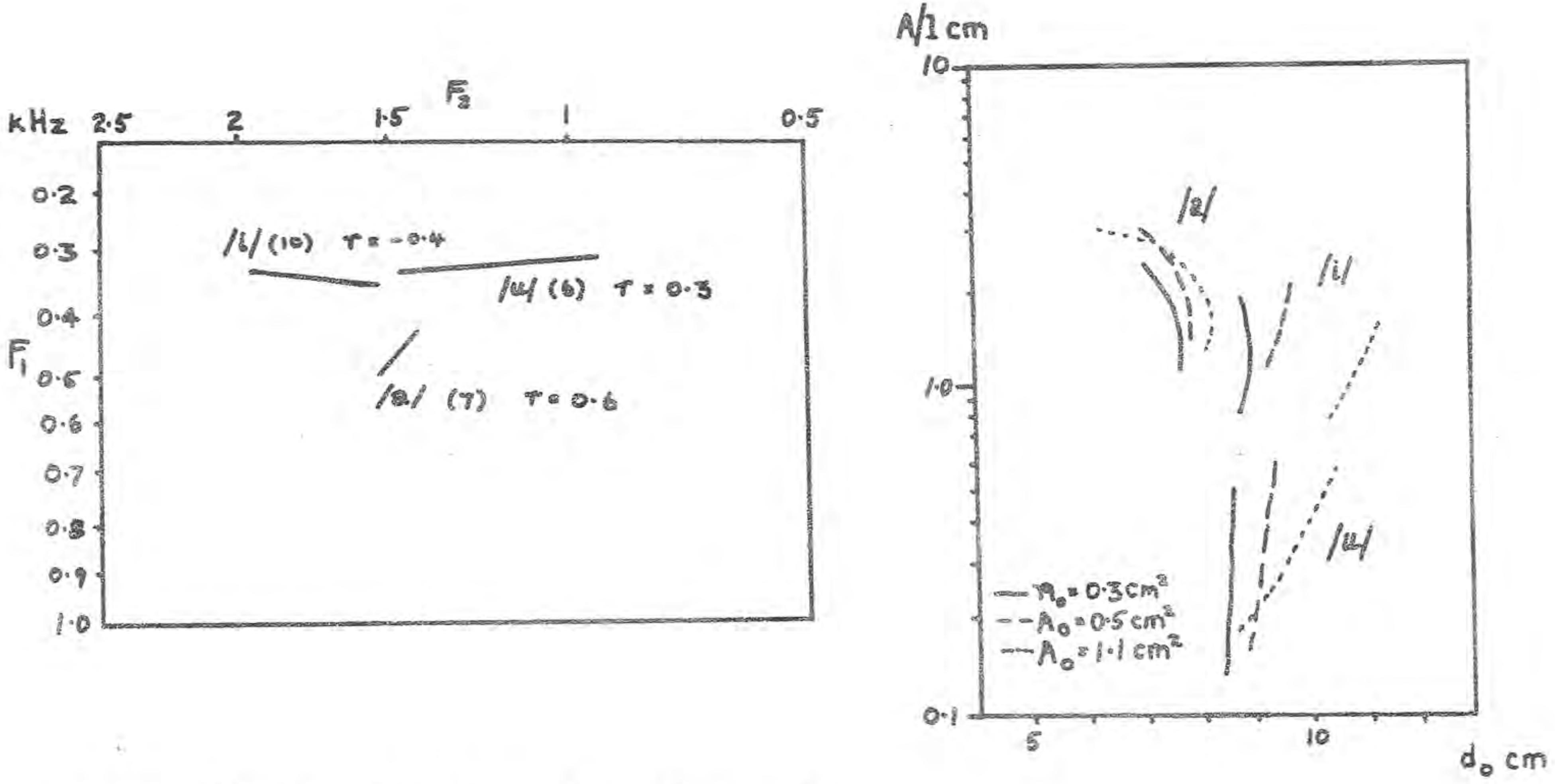

Figure 8. (a) The correlation between F_{1o} and F_{2o} in vowels (excluding labial, lateral, palatal and pharyngeal environments), and the regression of F_{1o} on F_{2o} for these vowels. (b) The articulatory contours of the spectra along the regression lines.

(c) Fig. 8b shows the articulatory contours for spectra along the regression lines, for three degrees of constriction. Considering each parameter independently, we can investigate articulatory centralization corresponding to the spectral centralization along the regression lines.

(i) The main modification to the resonator configurations concerns the mouth-opening parameter A/1. For all three phonemes, at any degree of constriction quoted, the centralization of spectra along the regression lines was associated with clear centralization of the mouth-opening to less extreme positions, falling to smaller than usual for /i/ and /a/ (to about A/1 = 1.0 cm., possibly related to failure to open the mandible) and rising to wider than usual for /u/ (to about A/1 = 0.7 cm., possibly related to failure to round the lips).

(ii) Little variation of the place of constriction would be necessary at each degree of constriction - except that there appears to be a slight centralizing movement of the place at the largest degree of constriction quoted (most pronounced for /a/).

(iii) All the spectra under consideration can be generated with $A_o \leq$ 1.1 cm^2, and the main effect of varying the degree up to this size is to shift the constrictions of all phonemes higher up the vocal tract. In addition, as already observed, there is the appearance of some slight centralization of the place of constriction at larger degrees of constriction. But this apparent centralization is false since the range of constriction locations is larger when A_o is large ($6 < d_o < 11$ cm. for $A_o = 1.1$ cm^2) than when it is small ($7 < d_o < 9$ cm. for $A_o = 0.3$ cm^2).

The regression lines mark only the beginning of vowel reduction - they cover the stressed vowel areas and the distinctive parts of the weak vowel areas (Fig. 2 (b, c). If the /i/ and /a/ regression lines are produced towards a spectrum of 350, 1250 Hz [the centre of the overlapping area on

Fig. 2c, discussed at § 4.2.2(a)], the three-parameter model indicates even further centralization of the mouth-opening around A/l = 0.7 cm. and now also a definite centralizing shift of constriction locations towards the velum.

(d) The following reconstruction of the articulatory correlates of spectral reduction is tentatively proposed:

(i) For the beginning of the spectral centralization, from the periphery and half-way in to the centre, the movement of the lips and mandible are gradually restricted to the middle range of mouth openings. A_o remains small.

(ii) For the remainder of the path to the centre with complete loss of spectral distinctions, there is continued narrowing of the range of mouth-openings. In addition, there is either (i) some gradual centralization of the constriction location towards the velum while the cross-section area remains small, or, (ii) the cross-section area increases and the constrictions are scattered all along the vocal tract as the configuration approaches the uniform tube. Intuitively, I prefer the former alternative. This can only be speculation, but the matter could be settled by cine-radiographic investigation of successive vocal tract states during the vowels of continuous speech.

4.3 A pharyngeal vowel allophone rule

4.3.1 Articulations

(a) Non-pharyngeal environments

/a/ - a pharyngeal constriction and large mouth-opening. In some non-pharyngeal environments, /a/ spectra have been found requiring velar or slightly palatal constrictions. This [æ] allophone will be disregarded now since the investigation has not aimed at describing allophonic alternation within the non-pharyngeal environments. The basic constrictions of /a/ are below the velum, i.e. pharyngeal.

/i/ – a palatal constriction and moderate or larger mouth-opening.

/u/ – a velar constriction and a small mouth-opening.

(b) Pharyngeal environments

/a/ – a pharyngeal (especially low pharyngeal) constriction and large mouth-opening.

/i/ – a pharyngeal (especially uvular) constriction and moderate or larger mouth-opening.

/u/ – a pharyngeal (especially uvular) constriction and small mouth-opening.

4.3.2 A feature framework based on the three-parameter model

(a) Traditionally, vowel articulations are described according to the position of the dorsal hump in a quadrilateral relative to the roof of the mouth, ignoring pharyngeal constrictions. This is only indirectly, and not always predictably, related to the actual configuration of the vocal cavities (cf. Fant 1960: § 2.33). In particular, it is difficult to fit "uvularization" into a framework that does not otherwise take into account the pharyngeal cavity. However, vowel articulations can be handled in terms of the three-parameter model, as has been done in preceding sections of this paper. This provides better correspondence between the articulatory and acoustic levels of description, and takes full account of the pharynx. An added advantage of this approach is that the same place features are used for both vowels and consonants. Such a scheme will be described in greater detail in a forthcoming article, but its application to the W.G. Eskimo allophones will be outlined here.

(b) Place features. The vocal tract can be divided into three main regions with regard to the speech functions of its vowel output - the pharyngeal part, the velar part and the palatal part. These regions correspond to the inferred places of articulation of the three vowel phonemes. Radiographic evidence for the same tripartite division of the vocal tract has been given by Lindblom and Sundberg (1969, 1971), who found three basic families of tongue articulations for Swedish vowels. Two features, palatal and pharyngeal, can be combined in a binary system to denote the three regions:

+palatal	-palatal -pharyngeal	+pharyngeal

For the allophones in the pharyngeal environments, it is necessary to subdivide the pharyngeal cavity into an upper part (uvular, for /i/ and /u/) and a lower part (below the uvula, for /a/), according to § 4.3.1(b). The feature uvular makes this division.

(c) Mouth-opening features. A distinguishing feature of /u/ is the small mouth-opening ratio A/l associated with lip-rounding, while non-labial /i/ and /a/ have larger ratios, the dividing value being approximately A/l = 0.7 cm. (cf. Fig. 5a). This difference will be denoted labial. Further, the range of mouth-openings was larger for /a/ than for /i/, corresponding to the traditional features of "open" and "close", or "low" and "high". This difference will be denoted open. Combining the two features labial and open, three basic degrees of mouth-opening are obtained:

+labial	-labial -open	+open

4.3.3 Distinctive features of W.G. vowels

(a) Non-pharyngeal environments. A minimum specification of distinctions is as follows: there are three vowel phonemes, one palatal (/i/) and two non-palatal (one labial, /u/, and one non-labial, /a/). This gives the following scheme (/i/ being redundantly non-labial):

pal	+ lab	− lab
+		i
−	u	a

(I)

Fig. 9a shows the spectra that can be generated within the limits of these distinctions. Line A–A' divides the spectral space into an area generated with palatal constrictions and an area generated with non-palatal constrictions. Line B–B' divides the spectral space into an area generated with lip-rounding and an area without lip-rounding. These lines have been plotted as follows:

(i) Line A–A' represents the spectra for d_o = 10 cms. (the palatal/velar boundary) and the full range of mouth-openings. The degree of constriction is taken to be a function of mouth-opening in so far as the latter is produced by jaw movements. The cross-section area A_o has therefore been increased in step with A/l, from A/l = 0.1 cm. and A_o = 0.3 cm^2 at A to A/l = 20 cms. and A_o = 3.1 cm^2 at A'.

(ii) Line B–B' represents the spectra for A/l = 0.7 cm. (the inferred boundary between labial and non-labial mouth-openings, Figs. 5a and 8b) and all places of constriction ($4 < d_o < 13$ cms.). Degrees of constriction have been chosen that are similar to the Russian vowel data quoted in § 2.5 from Fant (1960) – 0.5 cm^2 for a close palatal vowel and 1.1 cm^2 for a close non-palatal vowel (i.e. r_o values 0.4 cm. and 0.6 cm. respectively on Stevens and House's charts).

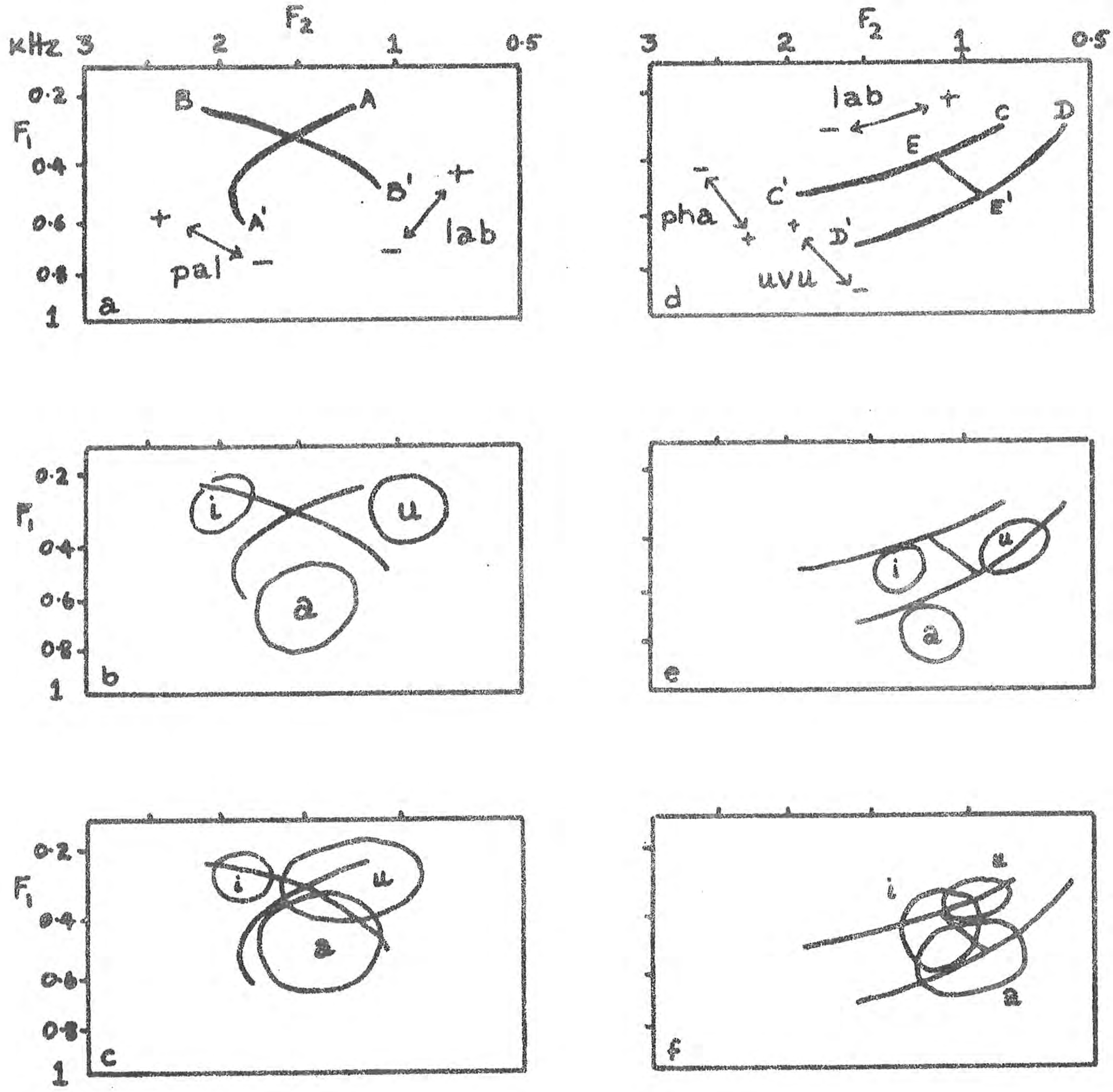

Figure 9. Acoustic correlates of distinctive features (a, d) compared with spectra observed in stressed vowels in carefully pronounced words (b, e) and continuous speech (c, f) in non-pharyngeal (a, b, c) and pharyngeal (d, e, f) environments.

Fig. 9 (b, c) shows the spectral areas of stressed vowels from Fig. 2 (a, b) superimposed on these contrasting areas.

(b) Pharyngeal environments. In the pharyngeal environments, distinctions are maintained by contrasting a low-pharyngeal (non-uvular) vowel (/a/) and two uvular vowels (one labial, /u/, and one non-labial, /i/). Fig. 9d shows the spectra that can be generated within the limits set by these distinctions. The line C-C' marks the pharyngo/velar boundary. The line D-D' divides the pharyngeal spectra into an area generated with uvular constrictions and an area generated with constrictions below the uvula. The line E-E' divides the spectral space into an area generated with lip-rounding and an area generated without lip-rounding. These lines have been plotted as follows:

(i) Line C-C' represents the spectra for d_o = 8 cms. and a small range of mouth-openings (as inferred from the informant's spectra). The degree of constriction has been set very small with small mouth-openings (A_o = 0.3 cm^2 and A/l = 0.1 cm. at C), and slightly larger with non-labial mouth-openings (A_o = 1.1 cm^2 for A/l > 0.7 cm.). A/l = 10 cms. at C'.

(ii) Line D-D' represents the spectra for d_o = 6 cms. The other two parameters, A/l and A_o, have been set as for C-C'.

(iii) Line E-E' represents the spectra for A/l = 0.7 cm. in the pharyngeal area. A_o has been set at 0.5 cm^2 (it was inferred from the informant's spectra that A_o was probably smaller in pharyngeal environments).

Fig. 9 (e,f) shows the spectral areas of stressed vowels from Fig. 2 (d, e) superimposed on the contrasting areas.

(c) Full feature specifications. A complete redundant specification, according to § 4.3.1 (a) and using the features outlined in § 4.3.2, is as follows:

	/a/	/i/	/u/
palatal	−	+	−
labial	−	−	+
pharyngeal	+	−	−
open	+	−	−

(II)

and in pharyngeal environments according to § 4.3.1 (b):

	/a/	/i/	/u/
palatal	−	−	−
labial	−	−	+
pharyngeal	+	+	+
open	+	−	−
uvular	−	+	+

(III)

4.3.4 The pharyngeal allophone rule

This rule transforms the feature matrix (II) into (III) before pharyngeal consonants.

(a) Vowels assume the same place of constriction as the following consonant:

$$\begin{bmatrix} +\text{syl} \\ -\text{cns} \end{bmatrix} \longrightarrow [+\text{pha}] \;/\; \_\_\_ \begin{bmatrix} +\text{cns} \\ +\text{pha} \end{bmatrix} \qquad \text{(IV)}$$

(b) The place of articulation in (IV) must be adjusted to give the exact location of the constriction – at the uvula for /i/ and /u/, below the uvula for /a/. A suitable solution would be an alpha-rule, where the sign of the feature [uvular] is determined by some characteristic distinguishing /a/ from /i, u/. There are several possibilities:

(i) /a/ is neither palatal nor labial, while /i/ and /u/ are either palatal or labial. A rule based on this difference is tricky to formulate, but one possibility is the following unorthodox adaptation of the alpha-environment convention (Harms 1968:71):

$$\alpha \left\langle \begin{bmatrix} -\text{pal} \\ -\text{lab} \end{bmatrix} \right\rangle \longrightarrow [-\alpha\ \text{uvu}] \qquad \text{(V)}$$

(ii) /a/ is redundantly [+pharyngeal] prior to the transformation, cf. (II) above, while /i/ and /u/ are [-pharyngeal]. This gives

$$[\alpha \text{pha}] \longrightarrow [-\alpha \text{uvu}] \qquad \text{(VI)}$$

(iii) /a/ is redundantly [+open], cf. (II) and (III) above, while /i/ and /u/ are [-open]. This gives:

$$[\alpha \text{opn}] \longrightarrow [-\alpha \text{uvu}] \qquad \text{(VII)}$$

Solutions (VI) and (VII) are much simpler than (V), nor do they require any new convention for expansion. Which of (VI) and (VII) is to be preferred? It is possible that there is a universal relationship between the openness of a back vowel and the lowness of its constriction in the pharynx. If so, a rule based on (VII) would not only generate the desired output but would also express a general phonetic fact. For a complete rule, (IV) and (VII) have been combined:

$$\begin{bmatrix} +\text{syl} \\ -\text{cns} \\ \alpha \text{opn} \end{bmatrix} \longrightarrow \begin{bmatrix} +\text{pha} \\ -\alpha \text{uvu} \end{bmatrix} \Big/ \ \underline{\qquad} \begin{bmatrix} +\text{cns} \\ +\text{pha} \end{bmatrix} \qquad \text{(VIII)}$$

4.3.5 Consonant adjustment

It has previously been inferred that the pharyngeal vowel allophones probably had fairly small degrees of constriction. This means that the tongue bulge would be very close to the posterior pharyngeal wall. The constriction for /a/ is lower than that for /i/ or /u/. It is a reasonable question to wonder whether (i) the consonant constriction following /a/ will also occur low in the pharynx opposite the bulge (a tongue movement of a few millimetres, possibly a continuation of the same muscular effort that is depressing and retracting the tongue for the vowel), or (ii) whether the consonant constriction following /a/ would be higher up, in the uvular region, according to the standard description of these consonants as uvulars, (meaning more complicated muscular activity involving relaxation of the low vowel constriction andcreation of a new bulge a few centimetres higher up). Assimilation of the consonant to the place of articulation of the vowel, according to the first alternative, is not improbable. The very grave quality of the informant's pharyngeal consonants has already been remarked on in this paper. Some further indication is given by the inferred constriction locations for formant transition terminal spectra in the following words:

word	syllable	F_{1t}, F_{2t}	inferred d_o
aqigsseq	aq-	800, 1200 Hz	5 cms
aqo	aq-	625, 1000 Hz	6 cms
arqa	arq-	750, 1150 Hz	5 cms
qajaq	-jaq	750, 1125 Hz	5 cms
ikeq	-keq	575, 1075 Hz	6.5 cms
eqeq	eq-	500, 1000 Hz	7 cms
	-qeq	500, 1000 Hz	7 cms
neqe	neq-	450, 1050 Hz	7.5 cms.

A very small degree of constriction has been assumed, the constriction being minimal the moment before full consonantal obstruction. The inferred consonant locations are lower after /a/ (d_o 5-6 cms.) than after /i/ (d_o 6-7 cms.). A <u>pharyngeal consonant adjustment rule</u> assimilating the consonant constriction to the vowel constriction, might be as follows:

$$\begin{bmatrix} +\text{cns} \\ +\text{pha} \end{bmatrix} \longrightarrow [\alpha\text{uvu}] \,/\, \begin{bmatrix} +\text{pha} \\ \alpha\text{uvu} \end{bmatrix} \text{———} \qquad \text{(IX)}$$

giving lower consonant constrictions after /a/. This is the extreme opposite of the rival view, that the /k-q/ opposition is a palatal vs. velar contrast rather than velar vs. uvular. The above interpretation of the acoustic data not only suggests that /q/ and /ʁ/ are sub-velar, but also that the place of articulation may occur anywhere in the pharyngeal cavity - at or below the uvula - depending on the preceding vowel. The evidence for this may seem slender, being based on a small sample of speech from one informant, but the point could be settled by reference to radiographs.

Bibliographical references

Chiba T. and Kajiyama M. 1941. The vowel, its nature and structure. Tokyo

Fant C.G. 1960. Acoustic theory of speech production. The Hague

Fant C.G. 1968. Analysis and synthesis of speech processes. In Manual of Phonetics. Ed. Malberg, the Hague

Harms R.T. 1968. Introduction to phonological theory. New Jersey

Hill A. 1968. Introduction to linguistic structure. New York

Lawrenson A.C. 1934. Two specimens of Eskimo. Maître Phonétique

Lindblom B. and Sundberg J. 1969. En fonetisk beskrivning av svenska vokaler. Pilus 1969, no. 2. Stockholm

Lindblom B. and Sundberg J. 1971. Acoustical consequences of lip, tongue, jaw and larynx movement. Pilus 1971, no. 2. Stockholm

Shearme and Holmes. 1962. An experimental study of the classification of sounds in continuous speech according to their distribution in the F1/F2 plane. Proceedings of the 4th International Congress of Phonetic Sciences. Helsinki

Stevens K. and House A.S. 1955. Development of a quantitative description of vowel articulation. JASA 27, 484–495

Stevens K. and House A.S. 1961. An acoustical theory of vowel production and some of its implications. Journal of Speech and Hearing Research 4, 303–320