



הטכניון – מכון טכנולוגי לישראל
Technion – Israel Institute of Technology

ספריות הטכניון
The Technion Libraries

בית הספר ללימודי מוסמכים ע"ש ארווין וג'ואן ג'ייקובס
Irwin and Joan Jacobs Graduate School

©

All rights reserved

*This work, in whole or in part, may not be copied (in any media), printed, translated, stored in a retrieval system, transmitted via the internet or other electronic means, except for "fair use" of brief quotations for academic instruction, criticism, or research purposes only.
Commercial use of this material is completely prohibited.*

©

כל הזכויות שמורות

אין להעתיק (במדיה כלשהי), להדפיס, לתרגם, לאחסן במאגר מידע, להפיץ באינטרנט, חיבור זה או כל חלק ממנו, למעט "שימוש הוגן" בקטעים קצרים מן החיבור למטרות לימוד, הוראה, ביקורת או מחקר. שימוש מסחרי בחומר הכלול בחיבור זה אסור בהחלט.

צמצום ספרתי של אוחות דבור
באמצעות התמרה פוריה הדיסקרטית

*

חבור על פרויקט

לשם מילוי חלקי של הדרישות לקבלת החוואר
מגיסטר למדעים
בהנדסת חשמל

מ א ח

יעקב שולמן



חוגש לסנסט של הטכניון - מכון טכנולוגי לישראל

ספטמבר 1974

ח י פ ה

חשרי חשל"ה

המחקר נעשה בהנחייתו של ד"ר ד. מלאך
מהפקולטה להנדסת חשמל - הטכניון, מכון טכנולוגי לישראל.
תודתי העמוקה נתונה לו על עזרתו הרבה.

תודתי נתונה גם לאנשים הבאים:
א. גוטמן אשר דרבן אותי לנצל נושא זה לעבודתי
לקראת התואר.

ד"ר מ. ביטרן על שיחותיו ועצותיו במשך עבודתי.

כמו כן ברצוני להודות להנהלת מפעל אלג'ים על
אשר הסכימה לממן מחקר זה שלולא כן לא היה
יוצא לפועל.

1	1. חקציר
3	2. רשימה סמליט וקצורים
4	3. מבוא
5	4. צמצום אותות דיבור
5	4.1 הדיבור
11	4.2 צמצום אותות דבור בשיטות הפרדה - הרכבה (analysis-synthesis)
14	4.3 איכות דיבור משוחזר
17	5. צמצום אותות דיבור באמצעות התמרת פוריה
17	5.1 התמרת פוריה הדיסקרטית וחכונותיה
19	5.2 התמרת פוריה מהירה
19	5.3 שמוש בהתמרת פוריה לצמצום אותות דיבור
24	5.4 צמצום אותות דיבור באמצעות התמרות אורתוגונליות שונות
25	6. תאור העבודה ותוצאותיה
25	6.1 הדימוי במחשב
27	6.2 בדיקות מובנות
29	6.3 תלות המובנות בשלושה פרמטרים בטיסיים
35	6.4 השפעת חלון Hann
37	6.5 קצב ההעברה ותלות המובנות בו
39	6.6 השפעת שנויים שונים בשיטת הצמצום
44	6.7 איכות הדיבור המשוחזר
44	6.8 חוסר התלות בין מובנות הדיבור המשוחזר לקרוסקורלציה ושגיאה ריבועית ממוצעת
47	7. סיכום והמלצות
47	7.1 סיכום האפשרויות האופטימליות לנצול שיטת השיאים המקומיים
48	7.2 המלצות להמשך
49	8. נ ס פ ח י ס
49	נספח א': פרטי הנסויים ותוצאות נסויי המובנות
56	נספח ב': לוחות המליט אשר שמשו לנסויים
57	נספח ג': ריכוז תוצאות שונות
	1.ג אחוז הקוים הנבחרים למי אזורים
	2.ג קצב העברה בנסויים שונים
	3.ג קרוסקורלציה ושגיאה רבועית ממוצעת
63	נספח ד': תכניות מחשב
81	9. רשימת מקורות

1. ת ק צ י ר

באפיקי תקשורת דבור מועברים השדרים באמצעות אות הדבור. קיבול המידע של אפיק הדרוש להעברת האות הוא גבוה בכמה סדרי גודל מהקצב בו האדם מסוגל להפיק או לקלוט מידע. צמצום תכולת אות הדבור יקטין את קבול האפיק הנדרש.

נערכו עבודות רבות בנושא של צמצום אות הדבור בשיטה של נתוח האות, צמצום ושחזור אות הדבור. חלקן מבוססות על ידיעת תהליך יצור הדבור ע"י האדם ומדידת הפרמטרים הקשורים לכך (פונקצית הערור והמסנן), חלקן האחר מבוסס על צורת הספקטרום של אות הדבור ומדידת הפרמטרים הקשורים בו.

מטרת פרויקט זה היתה בדיקת אפשרות צמצום תכולת אות דבור באמצעות התמרת פוריה הדיסקרטית של האות ומבלי להזדקק למדידת הפרמטרים של פונקצית הערור של קולות הדבור. האוזן רגישה מאד לשנויים קטנים ביותר בפונקצית הערור ולכן הסתמכות על פרמטרים אחרים מקנה יתרון לשיטה המוצעת.

שיטת השיאים המקומיים המוצעת מבוססת על נתוח אות הדבור ע"י התמרת פוריה הדיסקרטית וצמצום תכולתו ע"י בחירה של חלק מרכיבים אלה ואפוס כל האחרים. הבחירה מבוססת על צורת הספקטרום הדיסקרטי של אות הדבור היות והרכיבים הנבחרים הם אלה המהוים שיאים מקומיים הגדולים ביותר של הספקטרום הדיסקרטי. שחזור הדבור מן הרכיבים הנבחרים מבוצע ע"י התמרת פוריה דיסקרטית הפוכה. נתוני הרכיבים הנבחרים מועברים למשחזר בצורה של משרעת החלק הממשי, משרעת החלק המדומה והמקום הספקטרלי לגבי כל רכיב. המשרעת מועברת לאחר בצוע קוונטיזציה לוגריתמית.

המחקר בוצע ע"י דמוי במחשב ספרתי, תוך שמוש באלגוריתם של התמרת פוריה המהירה, נבדקה מובנות הדבור המשוחזר בתלות בשלושה פרמטרים בסיסיים: אורך הבלוק (מס. הדגימות) כבצוע ההתמרה הדיסקרטית, מספר הרכיבים הנבחרים ומספר הסביות בקוונטיזציה הלוגריתמית של משרעת הרכיבים הנבחרים. המובנות חושבה מתוצאות בדיקות שמיעה של 15 מלים עבריות בנוח שתי הברות כל אחת. הבדיקות נערכו עם 10 מאזינים בכל נסוי, ס"ה נערכו 30 נסויים. כקריטריון לווי שמשה טבעיות הדבור המשוחזר. המטרה היתה כמובן להשיג צמצום מירבי תוך פגיעה מזערית כמובנות ובטבעיות.

מסקנת המחקר היא שנתן להשיג דבור עם מובנות הברות של 75 אחוז ומעלה, עם קצב העברה של 3750 סביות לשניה (bits/sec). יחס הצמצום בדרך זו הוא 1:15. הטבעיות של הדבור המשוחזר בתנאים אלה ניתנה לדרוג כמספקת עד טובה. אורך הבלוק האופטימלי נמצא כ-16 מילישניות (128 דגימות), כאשר מותרים $1/32$ מרכיבי ההתמרה והקוונטיזציה מבוצעת עם 4 סביות בלבד.

שיטת שיאים מקומיים מצטיינת בירתון נוסף והוא חמסורח ספרתית של האוח המצומצם.
אפשר לנצל שיטת שיאים מקומיים להעברה בקצבי העברה סטנדרטיים של 9600, 4800 או 2400
טביות לשניה.

חוך כדי בצוע המחקר נעשה נסיון למצא התאמה בין המובנות כפי שנמדדה לבין
קרוסקורלציה או שגיאה רבועית ממוצעת שבין אות הדבור המצומצם לאות הדבור המקורי.
התאמה כזו לא נמצאה.

בסיום הפרויקט מובאות המלצות להמשך עבודה במטרה לשפר את החוצאות.

2. רשימת הסמלים והקצורים

- a מספר הסביות הנחוץ לתאור אורך הבלוק $N = 2^a$
- $A(1)$ הרכיב ה-1 של התמרת פוריה הדיסקרטית
- BW רוחב סרט
- DFT התמרת פוריה הדיסקרטית (Discrete Fourier Transform)
- FFT אלגוריתם מהיר לבצוע DFT (Fast Fourier Transform)
- f_s חדר הדגימה של אות הדבור
- f_c חדר הקטעון של מסנן
- Δf רזולוציה ספקטרלית
- I מובנות, נמדדת באחוזים
- $IDFT$ התמרת פוריה הדיסקרטית ההפוכה
- $IFFT$ התמרת פוריה מהירה הפוכה
- IPS אינץ' לשניה, יחידת מהירות הקלטה או השמעה בקלט-סרט
- N מספר הדגימות לבלוק
- n מספר רכיבי התמרת פוריה הנבחרים מתוך N רכיבי הבלוק
- r מספר סביות בקוונטיזציה של המסרעה
- $R_{xy}(C)$ פונקציית הקרוסקורלציה
- $S_A(1)$ המסרעה של הקו הספקטרי ה-1 של הספקטרום הדיסקרטי
- T_B משך הבלוק (sec)
- TR קצב העברת מידע, מבוטא במספר סביות לשניה (bits/sec)
- MSE שגיאה רבועית ממוצעת (Mean Square Error)
- ד"ב דציבלים (dB)
- ק"ה קילוהרץ (kHz)

3. מ ב ו א

באפיקי תקשורה דבור מועברים השדרים באמצעות אות הדבור. קיבול המידע של אפיק הדרוש להעברת האות הוא גבוה בכמה סדרי גודל מהקצב בו האדם מסוגל להפיק או לקלוט מידע.

צמצום תכולת אות הדבור יקטין את קבול האפיק הנדרש. השיטות לצמצום אות הדבור מתחלקות לשני סוגים עיקריים:

- א. שיטות סטטיסטיות לדחיסת חדר, משרעת וזמן [7].
- ב. שיטות הפרדה-הרכבה (analysis-synthesis) המבוססות על נתוח האות, צמצומו ושחזור הדבור ע"י הרכבה.

השיטות מהסוג השני מבוססות בחלקן על ידיעת תהליך ייצור הדבור ע"י האדם ומדידת הפרמטרים הקשורים בכך (פונקצית הערוור והמסנן), ובחלקן מבוססות על צורת הספקטרום של אות הדבור ומדידת הפרמטרים הקשורים בו.

צמצום תכולת אות הדבור פוגע בדרך כלל במובנות ובטבעיות הדבור בהשוואה למערכות רחבות טרט ובעלות רעש נמוך. היות ולא קיימת עד היום שיטה אוניברסלית אשר תצמצם את תכולת אות הדבור בצורה משביעת רצון, יש חשיבות לחקור שיטות נוספות להשגת מטרה זו. השיטה המוצעת כאן היא נתוח אות הדבור ע"י החמרת פוריה, צמצום תכולתו ע"י בחירה של חלק מתוך רכיבי ההחמרה ושחזור הדבור מרכיבים אלה ע"י החמרת פוריה הפוכה. שיטה זו מאפשרת גם חמסורת ספרחית של המידע, על כל יתרונוחיה.

4. צמצום אותות דבור

4.1 הדבור

השיטה הראשונית לתקשורת בין בני האדם היא בעזרת צלילי דבור. הדבור בצורתו האקוסטית מורכב מתנודות מהירות בלחץ האויר הנגרמות ע"י המנגנון הקולי. תנודות אלה נקלטות ע"י האוזן ומפוענחות במוח. התקשורת בדרך זו מוגבלת לטווחים קצרים בלבד. קיום תקשורת כזו לטווחים גדולים יותר מחייב הפיכת התנודות בלחץ האויר לאות חשמלי אשר אותו ניתן להעביר לטווחים גדולים בשיטות שונות. אוח חשמלי זה הוא האות בו יטופל בעבודה זו.

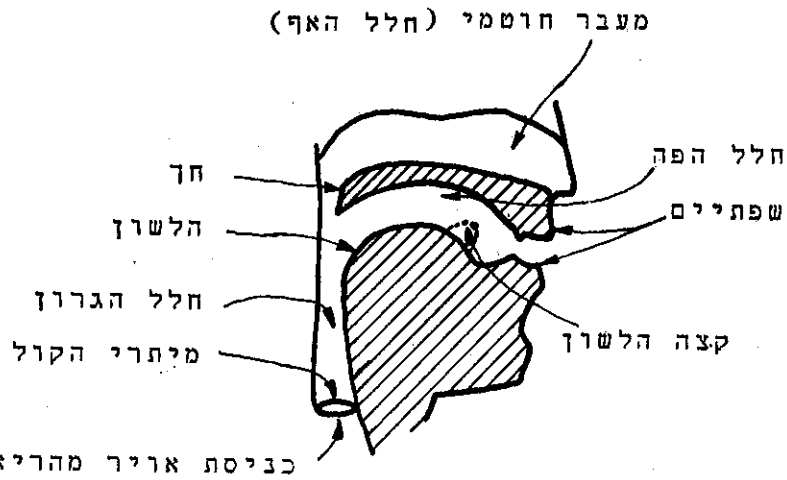
4.1.1 הפקת צלילי דבור ([8], [21], [22])

צלילי דבור מופקים ע"י אויר המשתחרר מן הריאות. זרם אויר זה מעורר את המעבר הקולי (vocal tract), המחואר סכמחית בצויר מס' 1, ויוצא דרך הפה או האף. המעבר הקולי הוא צינור אקוסטי אשר תחום בקצהו האחד במיתרי הקול ובקצהו השני בשפתיים. לצנור ראשי זה צמוד צנור משני אשר הוא המעבר החטמי (nasal tract). צנור משני זה מחובר או מחנתק מהצנור הראשי ע"י תנועת החך (velum). צורתו של המעבר הקולי נקבעת ע"י המקום והמצב של השפתיים, הלסת, הלשון והחך.

צלילי הדבור מופקים במערכת זו בשלושה אופנים בסיסיים. הצלילים הקוליים (voiced sounds) נוצרים ע"י ערוך המעבר הקולי בדפקים מחזוריים למחצה של זרמי האויר. דפקים אלה מקבלים ע"י מעבר זרם האויר מן הריאות דרך מיתרי הקול, אשר בתנודתם קוטעים לסירוגין את המעבר מן הריאות. תנודה זו של מיתרי הקול היא מאולצת ובדרך כלל בתדר הגבוה מחדר התנודות העצמיזח של המיתרים. צורתם של הדפקים היא משולשת והם בעלי מנת פעולה (היחס בין זמן הפתיחה לזמן המחזור) של 0.3 עד 0.7. החדר היסודי של רעידת מיתרי הקול, הנקרא גובה הצליל (pitch), נע בין 60 הרץ ל-500 הרץ. צלילי דבור המופקים באופן זה הם התנועות (vowels).

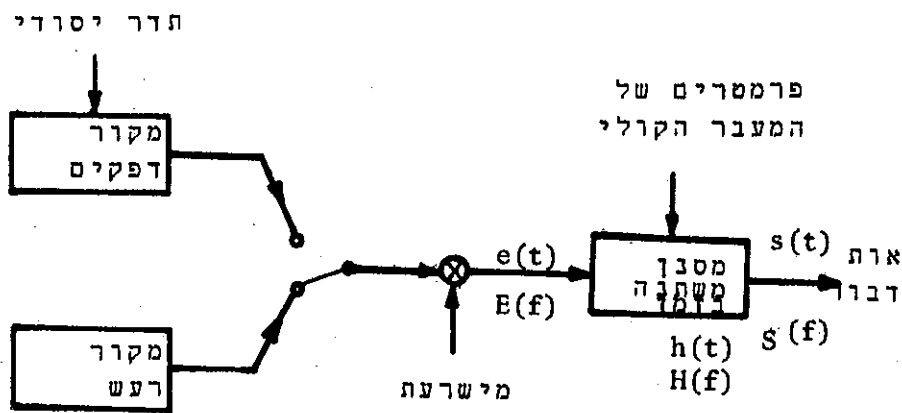
שני האופנים האחרים אחראים להפקת צלילים אטומים (unvoiced sounds). העצורים החוככיים (fricative consonants) מקורם במערבולות הנוצרות בעת זרימת האויר באלוצים צרים לאורך המעבר הקולי. מערבולות אלה יוצרות רעש רחב סרט, המעורר את המעבר הקולי.

המנגנון השלישי להפקת צלילי דבור מבוסס על סגירה מושלמת של המעבר הקולי, בנקודה כלשהי לארכו, ושחרור פתאומי של הלחץ הנוצר. כך נוצר דופק יחיד האחראי להפקת צלילים פוצצים (plosives) או סוחמים (stops).



ציור מס. 1 מונה סכמתי של המעבר הווקאלי אצל האדם

Fig. 1 Schematic drawing of the human vocal tract



ציור מס. 2 מודל לתאור תהליך הפקת הדבור

Fig. 2 Model for speech production

נוסף לצלילים הנוצרים ע"י כל אחד מהאופנים המוזכרים לעיל קימים צלילים המופקים ע"י שחוף שניים או אף שלושה מן האופנים.

כל מקורות הדבור הנ"ל הנם בעלי ספקטרום רחב סרט. המעבר הקולי מגיב לערור זה כמסנן בעל פרמטרים המשחנים בזמן. פונקציית התמסורת של מסנן זה מצטיינת במספר תדירויות תהודה טבעיות - פורמנטים (formants), לרוב שלוש מתחל ל-3000 הרץ. הפורמנטים מבטאים את תכונות התמסורת לצלילים של המעבר הקולי. מקורות הערור והתכנית של המעבר הקולי הנם בלתי תלויים (כמעט) ועל כן ניתן לתאר את תהליך הפקת הדבור כמתואר בציור מס' 2 [21]. בציור זה אות הדבור הוא המוצא של מסנן, משחנה בזמן, המחאר את תכונות התמסורת של המעבר הקולי. תוך כדי דבור רציף משנה המעבר הקולי את תבניתו, אך השנוי הוא איטי ועל כן תכונות תמסורת הצליל שלו משתנות באיטיות. בפרקי זמן של 10 עד 40 מילישניות, פרמטרים אלה קבועים למדי והם המגדירים את הפרמטרים של המסנן. עבור צלילים קוליים מעורר המסנן ע"י מקור דפקים מחזוריים למחצה, כאשר המרווח בין הדפקים הוא זמן המחזור של התדר היסודי. עבור צלילים אטומים המסנן מעורר ע"י מקור רעש בעל ספקטרום שטוח. בשני המקרים ישנה שליטה על המשרעת של האות הנכנס למסנן. אות המוצא $s(t)$ (הדבור) הוא הקונבולוציה של פונקציית הערור $e(t)$ עם תגובת ההלם של המסנן $h(t)$ (המעבר הקולי):

$$s(t) = \int_{-\infty}^t e(\tau) \cdot h(t-\tau) d\tau \quad (1)$$

4.1.2 הספקטרום של אות דבור

ספקטרום המשרעת של אות הוא, לפי הגדרה, הערך המוחלט של התמרת הפוריה שלו. אם נבצע התמרת פוריה של שני אגפי משוואה (1) ונקח את הערך המוחלט, נקבל:

$$|S(f)| = |E(f)| \cdot |H(f)| \quad (2)$$

כאשר:

$$S(f) - \text{היא התמרת פוריה של } s(t)$$

$$E(f) - \text{" " " " של } e(t)$$

$$H(f) - \text{" " " " של } h(t)$$

יש לזכור ששתי הפונקציות הערור והתמסורת הן משחנות בזמן ועל כן יש להחליף את הנתוח הספקטרי על פני אנטרוול זמן אינסופי בנתוח ספקטרי על פני אנטרוול סופי וקצר. לפי זה יהיה הספקטרום לזמן קצר בחלוי בזמן t וניתן ע"י המשוואה (Schroeder [22]):

$$|S(f,t)| = |E(f,t)| \cdot |H_t(f)| \quad (3)$$

כאשר $|H_t(f)|$ היא פונקציית התמסורת של המעבר הקולי בזמן t .

פרק הזמן של הניתוח הספקטרי, או "החלון הזמני", צריך להיות קטן ממשך צליל הדבור הקצר ביותר, כדי שאפשר יהיה להניח שפונקציית התמסורת הינה סטציונרית ומשוואה (3) תציין קשר מדויק. בכל מקרה אחר תציין משוואה (3) קרוב בלבד החלוי בסטציונריות $|H_t(f)|$. עבור קצב דבור רגיל "חלון זמן" אשר אינו עולה על 40 מילישניות מבטיח חנאים אלה.

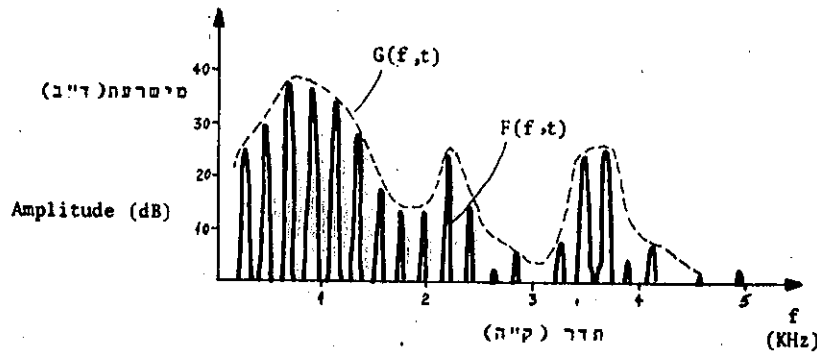
אם נרשום משוואה (3) בצורה לוגריתמית (בד"ב):

$$20 \log |S(f,t)| = 20 \log |E(f,t)| + 20 \log |H_t(f)| \quad (4)$$

נוכל לפרשה כאילו ספקטרום אות הדבור מורכב מהסכום של ספקטרום פונקציית הערוור ופונקציית התמסורת (בסקלה לוגריתמית כמובן).

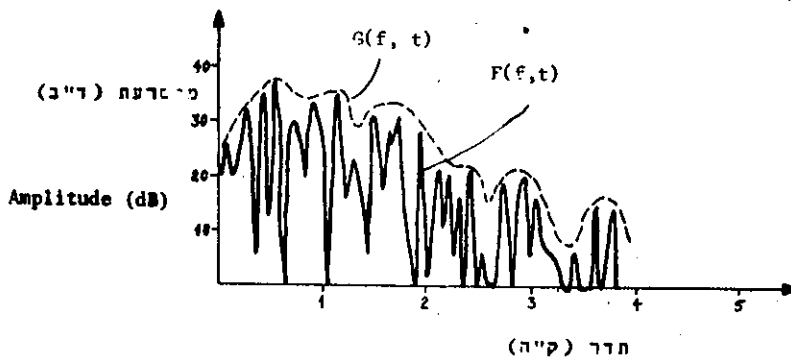
בציור מס' 3 מתואר ספקטרום צר זמן של צלילי דבור. ציור מס' 3א מתאים לצליל קולי והוא מאופיין ע"י קוים ספקטריים מוגדרים (חדים) אשר המרחק ביניהם (220 הרץ במקרה זה) הוא התדר היסודי ומעטפת בעלת שיאים רחבים אשר הם למעשה הפורמנטים. הקוים החדים מתארים את התדר היסודי וההרמוניות שלו. ציור מס' 3ב מתאים לצליל אטום והוא מצטיין בספקטרום רצוף (כמו רעש) בעל מעטפת הדומה בצורתה לקודמת. תאור זה של הספקטרום מוביל לתאור ספקטרום אות הדבור כמעטפת ספקטרלית $|G(f,t)|$ ומבנה ספקטרי עדין $|F(f,t)|$:

$$20 \log |S(f,t)| = 20 \log |F(f,t)| + 20 \log |G(f,t)| \quad (5)$$



a. voiced sound

א. צליל קולי



b. unvoiced sound

ב. צליל אטום

ציור מס' 3 ספקטרום האמפליטודה של צליל דבור

Fig. 3 Amplitude spectrum of speech sound

כאשר המבנה העדין $|F(f, t)|$ מצטיין בשיאים ספקטראליים מוגדרים עבור צלילים קוליים וכפונקציה רציפה בצלילים אטומים.

הפרוק של ספקטרום הדבור למרכיביו, קל יותר לפי התאור האחרון מאשר לפי התאור במשוואה (4). הסבה לכך נעוצה בקושי לשחזר את הספקטרום של פונקצית הערוור.

4.1.3 השמיעה

תהליך השמיעה איננו ברור במלואו ובמיוחד אין ידיעה מספקת על ההחמרה של ההזזה המכנית לפעילות עצבית המתבצעת באוזן הפנימית. הממברנה הבזילרית הנמצאת באוזן הפנימית אחראית להמרה זו והיא מתנהגת כמו קו חמסורת בלתי אחיד, עם רזולוציה ספקטראלית מוגבלת [22]. על כן מקובלת כיום ההנחה שההבחנה הספקטראלית המיוחסת לתהליך השמיעה נובעת מתהליך עבוד עצבי.

"חלון הזמן" עבור הממברנה הבזילרית יחסי הפוך לתדר והוא נע בין 0.2 מילישניות ל-12 מילישניות כלומר עבור מרבית צלילי דבור מחקבל, בעזרת הממברנה, ספקטרום עם רזולוציה גרועה. כך שרק התדר היסודי בצלילים קוליים מתגלה ע"י מנתח הספקטרום המכני (Flanagan [8] עמ' 126). כמו כן הרזולוציה הספקטראלית יורדת עם עליית התדר של אות הדבור [22].

רגישותה של האוזן לפאזה קטנה, במיוחד לשמיעה ממקור יחיד. עונת בפאזה יפגע מעט בטיב הדבור אך לא במובנחו. כל זה בתנאי שעות הפאזה איננו גורם להשהיה העולה על 50 מילישניות. אין זה בודאי נכון לאמר שהאוזן אדישה לפאזה.

תכונה נוספת היא הרגישות הגבוהה של האוזן למחזוריות של אות הדבור או במלים אחרות יכולתה של האוזן להבחין בתזוזה קטנה ביותר של התדר היסודי. לדוגמה בתדר של 1000 הרץ הובחנה סטיה של 3 הרץ. כפי שהוזכר קודם המנגנון של הבחנה זו איננו ברור די צרכו אך הוא מיוחס לעבוד בתהליך עצבי.

נסכם כאן את שלושת התכונות של השמיעה אשר הוזכרו לעיל:
א. האוזן מבצעת נחוח ספקטראלי צר זמן, עם רזולוציה היורדת עם עליית תדר אות דבור.

ב. בקליטה ממקור יחיד האוזן איננה רגישה לפאזה (בגבולות סבירים).

ג. האוזן רגישה בצורה יוצאת מן הכלל למחזוריות אותות דבור.

4.1.4 קיבול האפיק האנושי

נערכו נסויים שונים לקביעת הקצב המכסימלי בו מסוגל האדם לקלוט ולעבד מידע. בנסויים השונים נע הקצב בין 6 סביות לשניה ל-45 סביות לשניה ובכל מקרה אף נסיון לא הראה שהאדם מסוגל לעבד מידע בקצב העולה על 50 סביות לשניה [8, עמ' 7].

מנקודת מבט אחרת ננסה לראות מהו הקצב המכסימלי בו מסוגל האדם להפיק מידע בדבור. אם נחיס לפונמות (phonemes) כיחידות הקטנות ביותר המשמשות לתאור שדרי דבור הרי לפי Fant [7] כמות המידע המכסימלי בפונמה הוא 5 סביות לסימן (במקרה של שפה הכוללת 32 פונמות ובהנחה של הסתברות הופעה שווה לכל אחת מן הפונמות) בדבור מהיר מופקו כ-10 פונמות לשניה ועל כן הקצב המכסימלי של הזרמת מידע בשדרי דבור יהיה 50 סביות לשניה.

4.1.5 יתרות וצמצום אות דבור

ננסה להעריך את כמות המידע המועברת באפיק דבור רגיל או כמות המידע באות דבור לא מעובד. נסתכל על דבור כעל רוחב סרט מוגבל BW . במקרה זה לפי משפט ניקויסט (Nyquist) מוגדר אות זה במלואו ע"י $2BW$ דגימות בשניה. המשרעת של כל דגימה מוגדרת ע"י אחת מ- 2^x רמות, לאחר קוונטיזציה של המשרעת. כאשר r הינו מספר הסביות בקוונטיזציה של המשרעת, קצב העברת המידע TR יהיה במקרה זה

$$TR = 2BW \cdot r \quad [\text{bits/sec}] \quad (6)$$

בחקשורת בטיב מעולה מבוצעת קוונטיזציה ב-128 רמות (7 סביות), האות נדגם 8000 פעמים בשניה וקצב העברת המידע יהיה כ-56000 סביות לשניה. קצב זה מתאים לקיבול אפיק חקשורת דבור בעל רוחב סרט של 5 ק"ה ויחס אות לרעש של 30 ד"ב [7]. קצב זה גדול אם כך פי 1000 מקצב זרימת המידע שהאדם מסוגל להפיק או לקלוט. לכן לפי הערכה זו ישנה יתירות (redundancy) של 99.9% באות הדבור. המסקנה הנובעת מיד מכאן היא שניתן לצמצם בצורה ניכרת את אות הדבור וזאת מבלי לפגוע במידע המועבר.

קימות שיטות רבות לצמצום חכולת אות הדבור. ניתן לחלקן לשני סוגים עיקריים

לפי האופן בו מבוצע הצמצום:

- א. שיטות סטטיסטיות לדחיסת תדר, משרעת וזמן. בשיטות אלה מעוות האות בתהליך בלתי חוזר, ללא כל נסיון לעמוד על האופינים הפונטיים או הלשוניים. סקירת שיטות אלה ניתן למצא ב-Fant [7] ו-Flanagan [8]. הצמצום שנחן להשיג בשיטות אלה מגיע לחמישית לכל היותר.
- ב. שיטות הפרדה-הרכבה המבוססות על נחוח האות, צמצומו ע"י בחירת פרמטרים שונים המגדירים את תכונותיו החשובות ושחזור אות הדבור מתוך פרמטרים אלה בלבד.

למערכות אלה מקובל גם השם ווקודר (vocoder) על שם המערכת הראשונה מסוג זה של Dudley . בשיטות אלה משיגים כיום קצבי העברה שבין 500 ל-10000 סביות לשניה [9] , כלומר צמצום שבין פי 5 לפי 100.

צמצום אות דבור מבטיח תקשורת יעילה יותר באפיקים רועשים או בעלי קיבול מוגבל. כאשר היעילות מוגדרת במונחים של מובנות או איכות הדבור בחלות בקיבול האפיק נמצא בדרך כלל שמערכת המבוססת על צמצום הדבור עדיפה על מערכת רגילה בערכים נמוכים של קבול האפיק וההפך. מזה בערכים גבוהים.

היות וצמצום אות דבור מלווה בדרך כלל בהעברת אות מקודד הרי מתקבלת מאליו סודיות זסוימה (או פרטיות) בהעברת המידע וכמו כן ניתן בקלות להגדיל את מידת הסודיות ע"י קידוד אקראי.

יתרון נוסף לצמצום אות הדבור במקרה של צורך באכסון של מידע הדבור, היות ובמקרה זה ידרש קבול זכרון קטן יותר.

4.2 צמצום אותות דבור בשיטת הרכבה-הפרדה

העברת ערכי הפרמטרים השונים מן המפריד למרכיב נעשית בעזרת אותות אנלוגיים או ספרתיים. ההעברה מבוצעת בדרך כלל באפיק יחיד, מה שמצריך ריבוב האותות. במקרה של אותות אנלוגיים הריבוב יכל להעשות בדרך של חלוקת זמן או חלוקת תדר, והאות המרובב דורש יחב סרט מסוים להעברתו. היחס בין רוחב הסרט הדרוש להעברת אות הדבור המקורי לבין זה של האות המרובב מבטא את מדת הצמצום המושגת בווקודר הנדון. במקרה של אותות ספרתיים יבוצע ריבוב בחלוקת זמן (אם האותות אנלוגיים במקורם תבוצע המרה אנלוגית לספרתית תחילה) ונקבל ורם של דפקים בינריים המפורשים כסדרה של מספרים בינריים. זרם זה מוגדר בכמות הסביות לשניה שיש להעביר. להעברת דבור בעזרת אותות ספרתיים יש מספר יתרונות על העברה אנלוגית. הראשון הוא חסינות גבוהה בפני רעש וכמו כן הרעש, העוותים ושנויי השבח אָננם מצטברים בהעברה מרחקים גדולים, כפי שהדבר קורה עם אותות אנלוגיים. השני הוא האפשרות להסתיר ביעילות נ"י קידוד את המידע שבשדר כדי למנוע פענוח לא רצוי והבטחת תמסורת פרטית או אף סודית בהתאם לתחכום ההסתרה). יתרון שלישי נובע מכך שאותות ספרתיים עבור מידע שונה, כגון יבור, סלגרה, פסקימיליה או תמסורת נתונים נראים אותו דבר למערכת ההעברה וניתן לשלבם ללא ולחבנה. יתרון נוסף נובע מן האפשרות לשלב פונקציות של העברה ותחוג. עם זאת יש לזכור ויתרונות אלה מושגים בדרך כלל על חשבון רוחב הסרט הגדול הנדרש להעברת אותות ספרתיים.

נחאר בקצרה ארבעה סוגים ערוניים של ווקודרים:

- א. ווקודר ערוצי (channel vocoder) [7, 8, 13, 14, 22].
- ב. ווקודר קורלציה (correlation vocoder) [8, 22].
- ג. ווקודר פורמנטי (formant vocoder) [6, 7, 19].
- ד. ווקודר מעורר בדבור (voice excited vocoder) [7, 8, 13, 22].

א. ווקודר ערוצי

יחידת הפרדה של ווקודר ערוצי מורכבת משני חלקים: מנחה ספקטרום ומגדיר פונקציה הערור. מגדיר פונקציה הערור כולל גלאי המפריד את הדבור לקטעים קוליים וקטעים אטומים וכך מודד את התדר היסודי במשך כל קטע קולי. מנחה הספקטרום מודד בקרוב את המעטפת של ספקטרום הדבור.

מנחה זה כולל n ערוצים אשר כל אחד מהם כולל מסנן מעביר תחום, מישר ומסנן מעביר נמוכים. המסננים מעבירי תחום מסודרים כך שיכסו ברציפות את כל תחום הדרי הדבור המענין אותנו, המפריד מפיק אם כך $2 + n$ פרמטרים המועברים למרכיב. המרכיב כולל מערכת של מסננים מעבירי תחום המעוררים ע"י מקור רעש או מקור דפקים בהתאם לנתון אטום/קולי המגיע מן המפריד. תדר מקור הדפקים משתנה ונשלט ע"י התדר היסודי מן המפריד. תפוקת n המסננים מסוכמת, לאחר העברה דרך אפננים לקביעת המשרעת לפי ערכי n הפרמטרים של הספקטרום, וכך משוחזר הדבור. בנסויים השונים שנערכו נע מספר הערוצים בין 8 ל-18 והצמצום המושג מגיע עד כדי 20. רוחב הסרט הנדרש להעברת האותות בין המפריד למרכיב הוא בסדר גודל של 300 עד 400 הרץ. בהעברה ספרתית של פרמטרים אלה מגיעים לקצבי שדור של 2000 סיביות לשניה.

צמצום נוסף הושג ע"י Peterson [17] אשר הציע להעביר רק שלושה עד חמישה אותות הגדולים ביותר מבין 18 ערוצי הספקטרום של הווקודר, תוך הקפדה שלא יבחרו ערוצים סמוכים. בדרך זו ניתן לצמצם את קצב ההעברה פי 2 ביחס לווקודר ערוצי כלומר עד כדי 1000 סיביות לשניה.

למרות הצמצום הרב המושג בווקודר ערוצי אין הוא שמושי ביותר בגלל איכותו הירודה של הדבור המשוחזר המצטיין ב"צליל של מכונה" ובטבעיות פגומה. הסיבה לכך נעוצה כנראה בקושי להגדיר את פונקציה הערור ובמדודה גסה מדי של ספקטרום המעטפת. קים קושי רב להחליט בדיקנות האם הצליל קולי או אטום וקושי רב יותר למדוד בדיוק רב את התדר היסודי של הדבור ובמיוחד כאשר תחום התדרים הנמוכים של הדבור איננו קים כלל (מה שנכון במרבית המקרים כמו בערוץ טלפון). היות והאוזן רגישה מאד לשנויים קטנים בגובה הצליל סובל ווקודר זה מטיב לקוי.

ב. ווקודר קורלציה

פונקציית האוטוקורלציה של אות שזה להחמרת פוריה של ספקטרום ההספק צר-זמן של האות ולהפך. לכן המידע הספקטרי ניתן לתאור במונחים של פונקציית האוטוקורלציה. מכאן שאות הדבור יכול להיות מיוצג ע"י פונקציית אוטוקורלציה צר זמן במקום ע"י ספקטרום צר-זמן. זהו הבסיס לווקודר הקורלציה, אשר בו מחושבים רכיבן פונקציית האוטוקורלציה עבור השהיות שונות, וביחד עם נחוני פונקציית הערוך הינם מועברים למרכיב. שם משוחזר הדבור ע"י הכפלת תפוקת מקור הערוך ברכיבי פונקציית האוטוקורלציה וסיכום האוחות הבודדים על פני קו השהיה מתאים. רוחב הסרט הדרוש להעברת האות המצומצם הוא כ-400 הרץ. אך להעברת רכיבי האוטוקורלציה נדרש יחס אות לרעש טוב יותר (או יותר סביות לדגימה) מאשר להעברת רכיבי הספקטרום [22]. כמו כן הרזולוציה הספקטרלית האקויוולנטית, בלתי תלויה בתדר, דבר שמונע הרחבת רוחב הסרט של המסננים בתדרים הגבוהים יותר (כפי שהאוזן מתנהגת). מכל זה נובע שכדי להשיג בשיטה זו דבור באיכות הדומה למושגת בווקודר ערוצי יש צורך במספר רב יותר של ערוצים [22].

ג. ווקודר פורמנטי

הווקודר הפורמנטי מבוסס על תאור מעטפת ספקטרום צר-זמן של אות הדבור ע"י התדרים והמשרעות של הפורמנטים העקריים, ותאור המבנה העדין ע"י מקור רעש או דפקים, בהחלטת החלטה אטום/קולי ובמקרה של צליל קולי גם ע"י תדר הדפקים. ממוש המרכיב יכול להעשות בשני אופנים: מקבילי או טורי (קסקדה). הממוש המקבילי מאפשר שחזור טוב יותר של העצורים ואילו הממוש הטורי מאפשר שחזור טוב יותר של התנועות. באחד הווקודרים המקביליים מעבירים 8 פרמטרים סך הכל [8 עמ' 262]: תדר ומשרעת של כ"א משלושת הפורמנטים הראשונים, תדר ומשרעת של התדר היסודי. המרכיב כולל שלושה מסננים מעבירי תחום בעלי תדר מרכזי משתנה כ"א לפי תדר הפורמנט ומסנן רביעי מעביר נמוכים מתחת ל-400 הרץ. עצמת אות המוצא מכל מסנן נקבעת לפי משרעת הפורמנט המחאים, או המשרעת בתחום התדרים הנמוך לגבי המסנן של 400 הרץ. המסננים מעוררים ע"י מקור דפקים או מקור רעש כאשר החלפה מבוצעת לפי העצמה של הפורמנט השלישי ותפוקותיהם מסוכמות לשחזור אות הדבור. רוחב הסרט הנחוץ הוא 100 עד 200 הרץ, שהיא כמחצית רוחב הסרט הנחוץ לווקודרים ערוציים. עם זאת המובנות המושגת כיום בווקודר פורמנטי נופלת מזאת המושגת בווקודר ערוצי, כן ישנה בעיה של הגדרה מדויקת של תדרי הפורמנטים.

ד. ווקודר מעורר דבור

ווקודר מעורר בדבור בא להתגבר על הקושי בהגדרת הפרמטרים של פונקציית הערור. הרעיון הוא להמנע מהצורך להגדיר את פונקציית הערור ע"י העברת חחום התדרים התחתון של אות הדבור כפי שהוא מן המפריד למרכיב. המבנה של ווקודר ערוצי מעורר בדבור דומה במבנהו לווקודר שתואר לעיל אך במקום החלק המגדיר את פונקציית הערור ישנו פשוט מסנן מעביר נמוכים המעביר את החלק התחתון של תדרי הדבור ("base band"). אות זה נושא את נתוני פונקציית הערור והוא משמש במרכיב (לאחר עבוד נוסף של הרחבה ושיטוח הספקטרום) כמעורר של מערכת מסננים הזוהה לזו הקימה בווקודר ערוצי. איכות הדבור שמחקבל בשיטה זו אינה נופלת מאיכות הדבור המקורי, אך הוא דורש רוחב סרט גדול יהסיה והצמצום המשוג הוא ביחס של שלוש לאחד בלבד. קיימים ממושים מסוג זה גם עבור הסוגים האחרים של הווקודרים שהוזכרו כאן.

מקובל היום לערוך מחקר בתחום הדבור בעזרת דמוי במחשב ספרתי כך שאפשר למצא בספרות את החוצאות של דמוי כזה עבור סוגים שונים של ווקודרים. נזכיר כאן את Schaffer [21] ו-Larkin [14]. האחרון ביצע דימו של ווקודר ערוצי תוך שמוש בהחמרת פוריה מהירה (FFT) והשיג דבור באיכות גבוהה.

4.3 איכות הדבור

האיכות של דבור משוחזר נמדדת ע"י שלושה קריטריונים בסיסיים: מובנות, טבעיות (naturalness) והאפשרות לזהוי דובר. החשוב מבין שלושם הוא מובנות הדבור. האפשרות לזהות את הדובר רצויה אמנם אך היא שלישית בדרגת החשיבות ובמרבית המקרים לא נעשה כל נסיון להשיגה.

המובנות היא היכולת של המאזין להבין את אשר נאמר, והיא נמדדת באחוזים של יחידות הדבור הנקלטות נכון. הטבעיות נמדדת בסולם של חמש דרגות (אשר יוגדר בהמשך). אפשרות זהוי הדובר מוערכת לעת עתה רק במונחים של המספר היחסי של המאזינים אשר זיהו נכונה את האדם הדובר.

בהעברה של דבור רגיל (בלתי מעובד) יש כמובן קשר אחד-לאחד בין איכות הדבור בהתייחסות למובנות לבין טבעיות הצליל. לא כן הדבר בדבור המשוחזר בווקודר, שם המובנות יכולה להיות גבוהה למרות טבעיות נמוכה של הצליל. מצד שני סביר לצפות שאם הצליל נשמע טבעי מאד תהיה המובנות גבוהה וכן גם תהיה אפשרות לזהוי הדובר.

המובנות תלויה בדיוק ההעברה של ספקטרום הדבור. עם זאת אין צורך בהעברה מדויקת מאד של הספקטרום כדי להשיג מובנות גבוהה. נוסף לזה אין המובנות מושפעת מדיוק ההעברה של החדר היסודי או הגוון (1) (timbre) של הקול אך כן מושפעת מדיוק בהחלטה אטום/קולי. נסוי להערכת דבור משוחזר כולל הקראה של רשימת הברות, מלים או משפטים לקבוצה של מאזינים. אחוז היחידות הנקלטות נכון, מתוך סך הכל היחידות שהוקראו נקבע כמובנות. ההקראה וההאזנה נעשים ללא מגבלות זמן. להבטחת אובייקטיביות ההערכה של מובנות הדבור צריכה הרשימה של החומר להקראה לייצג את הסטטיסטיקה הצלילית של השפה. לצורך זה פותחו בשפה האנגלית רשימות של מלים לנסויים, המאוזנות מבחינה פונטית, כגון Harvard PB Lists או CNC Lists כאשר כל רשימה כוללת 50 מלים [15]. לא ידוע על קיום לוחות כאלה בשפה העברית. עם זאת המובנות הנמדדת כך איננה אבסולוטית היות והיא תלויה במספר גורמים:

- א. סוג יחידות הדבור המוקראות. יתקבלו מספרים שונים עבור משפטים, מלים לא קשורות או הברות סתמיות. לדוגמה בערוץ רועש, ביחס אות לרעש של 0 ד"ב המובנות עבור הברות סתמיות תהיה 40% בלבד. לעומת זאת המובנות של מלים בתוך משפטים תהיה 70% ואילו המובנות של ספרות תגיע ל-100% [7].
- ב. גודל המילון ממנו נבחרו השדרים ומידת ההכרה של המילון ע"י המאזינים. להדגמת חלות זו נפנה שוב לערוץ רועש בו מועברות הברות סתמיות. ביחס אחת לרעש של 0 ד"ב ובמקרה של טקסט הכולל הברות מתוך מילון בלתי מוגבל המובנות היא 40 אחוז. כאשר באותם תנאים מועבר טקסט של הברות הנבחרות מתוך 256 הברות אשר ידועות מראש למאזינים המובנות עולה ל-70 אחוז, ואם נצמצם את המילון ל-32 הברות בלבד תגיע המובנות ל-95 אחוז [8].
- ג. מידת האימוץ של המאזינים.

4.3.2 טבעיות הצליל

כיום הטבעיות של הצלילי דבור ניתנה להערכה רק ע"י מאזין מומחה ומאומן היטב. להלן חמש הדרגות בסולם הטבעיות [1]:

- א. מצוין - מתיחס להעדר כל צליל זר.
- ב. טוב - מורגשת נוכחות צלילים זרים אשר אינם פוגעים ברושם של דבור אמיתי.
- ג. מספיק - הדבור כולל צלילים זרים ועוותים אשר מפריעים אמנם לקליטת הדבור אך עם זאת כל צליל שומר על האיכות הנחוצה לקליטה נכונה.
- ד. גרוע - הצליל איננו דומה יותר למקורי אך בכל זאת ניתן להכיר בו ע"י קריטריונים מסוימים.
- ה. גרוע מאד - מצוין העדר כל רמז בצליל הנשמע לצלילי הדבור אשר הועברו.

(1) הגוון של הקול - הגוון מוגדר ע"י הרוחב של שיאי הפורמנטים וע"י השפוע של המעטפת הספקטרלית הממוצעת של צלילי הדבור.

טבעיות הצליל המתקבל תלויה מאד בנאמנות ההעברה של המעטפת הספקטרלית והחדר היסודי, ותלויה מעט בנאמנות העברת הגוון. לנאמנות ההעברה של המעטפת הספקטרלית השפעה גדולה יותר על הטבעיות מאשר על המובנות. עם זאת אי דיוק קטן בהעברת הגוון, במיוחד ברוחב הפורמנטים ובשפוע המעטפת הספקטרלית כמעט ואיננו פוגם בטבעיות. להגדלת טבעיות דבור משוחזר יש צורך לא רק בהבחנה מדויקת של אופי פונקציית הערור (רעש, דופק או שניהם גם יחד) אלא במקרה של צליל קולי גם בהעברת צורת הדופק של החדר היסודי.

4.3.3 אפשרות זהוי דובר

האפשרות לזהוי הדובר מושפעת מהדיוק בהעברת גוון הקול ועצמת החדר היסודי. לדיוק ההעברה בגדלו היחסי של החדר היסודי השפעה מכרעת על אפשרות זהוי דובר. מעניין לציין שסטיות מהיגוי נכון עוזרות לזהוי דובר ומשמשות אף בסיס לבצוע החלטה. האפשרות לזהוי דובר, לאחר עבוד אות הדבור בווקודר למשל, מוערכת לעת עתה רק ע"י המספר היחסי של המאזינים אשר זיהו נכונה את האדם הדובר.

5. צמצום אותות דבור באמצעות התמרת פוריה

5.1 התמרת פוריה הדיסקרטית ותכונותיה

התמרת פוריה הדיסקרטית (DFT) של הסדרה:

$$X(k), \quad k=0,1,\dots, N-1$$

המורכבת מ-N מספרים קומפלכסיים, מוגדרת כ:

$$A(1) = \frac{1}{N} \sum_{k=0}^{N-1} X(k) \exp(-2 \pi j k 1/N) \quad (7)$$

כאשר $A(1)$ הוא הרכיב ה-1 של DFT וגם הוא מספר קומפלכסי. לנוחיות הרישום מקובל לרשום זאת כך:

$$A(1) = \frac{1}{N} \sum_{k=0}^{N-1} X(k) W_N^{-1k} \quad (7')$$

היכן ש:

$$W_N = \exp(2 \pi j/N) \quad (8)$$

אם $X(k)$ הנם ערכי פונקציה בנקודות דיסקרטיות בזמן (דגימות) אזי 1 יבטא את החדר. באופן דומה אפשר לבטא את $X(k)$ כהתמרת פוריה הפוכה (IDFT) של $A(1)$, כלומר:

$$X(k) = \sum_{l=0}^{N-1} A(1) W_N^{1kl} \quad (9)$$

תכונות התמרת פוריה הדיסקרטית ניחן למצא אצל Cooley [5], ו-Gold [12]. הקשר וההשוואה להתמרת פוריה הרציפה ניחן למצא אצל Gentleman [11] ו-Cohran [4].

נציין פה רק שתי תכונות של התמרת פוריה הדיסקרטית.

$$A(1) = A^*(-1) \quad (10') \text{ ממשי גורר } X(k)$$

$$A(-1) = A(N-1) \quad (10'') \text{ מחזוריות } A(1) \text{ גוררת}$$

כאשר * מסמן ערך צמוד. נובע מכאן שכאשר הסדרה הזמנית ממשית הרי החלק הממשי של ההתמרת יהיה זוגי ואילו החלק המדומה יהיה בלתי זוגי. במלים אחרות נוכל לאמר שהחלקים הממשיים של רכיבי התמרת פוריה יהיו סימטריים סביב הרכיב $N/2$ והחלקים המדומים יהיו אנטיסימטריים סביב אותו רכיב.

כדי לבצע התמרת פוריה דיסקרטית של פונקציה זמנית רציפה יש לדגום תחילה את

הפונקציה בחדר fs ולבצע קוונטיזציה של עצמה. החדר fs חייב להיות לפחות כפליים

מהחדר הגבוה ביותר הכלול בפונקציה. סדרת הדגימות מחולקת לקבוצות עוקבות של N דגימות

ובכל פעם מבצעים התמרת פוריה של N דגימות. קבוצה כזו של N דגימות תקרא בלוק ומשך הזמן

הנחוץ לקבלת N הדגימות יקרא אורך הבלוק T_B והוא נתון ע"י:

$$T_B = N/fs \text{ (sec)} \quad (11)$$

ה-DFT מחמיר את הפונקציה הזמנית למשור החדר, כאשר הרזולוציה Δf נקבעת ע"י חדר הדגימה f_s ומספר הדגימות לבלוק:

$$\Delta f = f_s/N = 1/T_B \text{ (Hz)} \quad (12)$$

בהנחה שהפונקציה הזמנית ממשיה מספיק לדעת $N/2$ מהרכיבים במשור החדר היות ו- $N/2$ הנותרים יהיו ידועים לפי תכונות הסימטריה שבשיוון (10) לעיל.

נדגיש שוב שהרכיבים של החמרה פוריה הדיסקרטית הם קומפלכסיים כלומר:

$$A(1) = \text{Re} \{A(1)\} + j \text{Im} \{A(1)\} \quad (13)$$

ספקטרום המרעת של הפונקציה $X(k)$ יהיה נחון ע"י:

$$S_A(1) = (\text{Re} \{A(1)\})^2 + (\text{Im} \{A(1)\})^2 \quad (14)$$

כאשר $S_A(1)$ הוא הקו ה-1 של הספקטרום הדיסקרטי. ה-DFT מנוצל לעתים קרובות כקרום למציאת החמרה פוריה הרציפה. למרות שחלק גדול מהתכונות של החמרה פוריה הרציפה נכון גם עבור DFT אין היא מהוה קרוב טוב להחמרה הרציפה אלא אם קוימו תנאים מתאימים וגם אז חשוב לזכור ההבדלים ביניהם. החמרה פוריה הרציפה פועלת על פונקציה זמנית רציפה ובלתי מוגבלת בזמן. החמרה פוריה הדיסקרטית הינה פונקציה דיסקרטית מחזורית עם מחזור השוה לאורך הבלוק. מעטפת הדגימות במחזור אחד של הפונקציה הדיסקרטית זהה להחמרה פוריה הרציפה בחסום זה.

שתי חופעות גורמות לשגיאה בקרוב ההחמרה הרציפה ע"י DFT. התופעות הן קפול

(Aliasing, Foldover) וזליגה (leakage).

דגימת הפונקציה הזמנית חיבה כאמור להעשות בתדר שהוא לפחות כפליים מהתדר הגבוה ביותר הכלול בפונקציה. אי קיום תנאי זה יגרור שגיאה בחשוב ההחמרה הדיסקרטית עקב חופעת הקפול הגורמת לתדרים גבוהים (מעל מחצית תדר הדגימה) להראות כתדרים נמוכים. למניעת הקפול יש להגביל ראשית את רוחב הסרט של הפונקציה הזמנית בעזרת מסנן Antialiasing וכן לדגום בתדר גבוה ככל האפשר.

פעולת ההחמרה הדיסקרטית על קבוצה סופית של דגימות ניתנת לתאור כהכפלה של דגימות הפונקציה בחלון זמן מלבני. במשור החדר תגרור פעולה זו לקונבולוציה בין ההחמרה של הפונקציה הזמנית להחמרה של פונקצית החלון. היות וההחמרה של פונקצית החלון מצטיינת באונות צד (side lobes) יופיעו במשור החדר רכיבים פרזיטיים. תופעה זו נקראת זליגה. להקטנת תופעת הזליגה (במקרים בהם שגיאה זו מפריעה) משתמשים ב"חלונות זמן" בעלי צורה השונה ממלבנית (כגון משולשת, Hann, Hanning ועוד), והמצטיינים באונות צד נמוכות יותר.

5.2 התמרת פוריה מהירה

בצוץ החשוב של DFT של N דגימות ישירות על פי ההגדרה דורש N^2 הכפלות קומפלכסיות ו- N^2 (למעשה $N(N-1)$) חבורים קומפלכסיים. מסתבר שזו דרך בזבזנית מאוד לחשוב ההתמרה הדיסקרטית. התמרת פוריה מהירה (FFT) הינו אלגוריתם מחוכם לחשוב מהיר יותר של התמרת פוריה הדיסקרטית. בעזרת ה-FFT אפשר לחשב את ה-DFT של N נקודות בעזרת $2N \ln N$ הכפלות קומפלכסיות בלבד, מה שמביא לחסכון (עבור N מספיק גדול) בזמן המחשב הדרוש לבצוע החשוב. כמו כן מאפשר אלגוריתם זה בנית המרה לבצוע ההתמרה הדיסקרטית בזמן אמיחי.

היות והתמרת פוריה הדיסקרטית ההפוכה מוגדרת באותו אופן כמו ההתמרה הישרה הרי גם אלגוריתם ה-FFT פועל עליה ונקרא לו אז התמרת פוריה הפוכה מהירה (IFFT).

האור האלגוריתמים השונים של FFT ודרכי קבלתם ניתן למצא ב-Gold [12] Cohran [4], Gentleman [11] ו-Brigham [2].

5.3 שמוש בהתמרת פוריה לצמצום אותות דבור

5.3.1 שיטת שיאים מקומיים

שיטות הפרדה-הרכבה השונות שתוארו בטעיף 4.2 לעיל חסרונן הבולט הוא הצורך למדד את הפרמטרים של פונקצית הערור. היות והאוזן רגישה מאד לשנויים בפרמטרים אלה נדרש דיוק רב במדידתם ובחעדור דיוק נפגעת איכות הדבור המשותזר. השיטה המוצעת כאן איננה דורשת כל ידיעה של פונקציה הערור. צמצום תכולת אות הדבור מושג ע"י נתוח אות הדבור באמצעות התמרת פוריה הדיסקרטית (תוך ניצול אלגוריתם ה-FFT), בהירת הקוים הדומיננטיים של ספקטרום המשרעת, העברת רכיבי ההתמרה המתארים קוים אלה לאחר בצוץ קוונטיזציה לוגריחמית של משרעתם, אל המרכיב.

המפריד בציור מס' 4 כולל: מסנן מעביר נמוכים, ממיר אנלוגי לספרתי (A/D converter), מחמר FFT, בוחר קוים (בוחר שיאים מקומיים), מעגל קוונטיזציה לוגריחמית ומקודד.

חחילה מסונן האות ע"י מסנן מעביר תדרים נמוכים, בעל תדר קטעון f_c . חפקיד המסנן לסלק רכיבי אות, מעל התדר f_c , למניעת שגיאות בחשוב ההתמרת פוריה עקב תופעת הקפול (Folding or Aliasing) [12]. אות הדבור האנלוגי הופך לדסקרטי בזמן ובמשרעת תוך כדי מעבר דרך ממיר אנלוגי לספרתי. פעולת הממיר כוללת דגימת אות הדבור וקוונטיזציה של הדגימות לפי x_1 סביות (מספיקות 8, בדרך כלל, להבטחת דיבור בטיב מעולה).

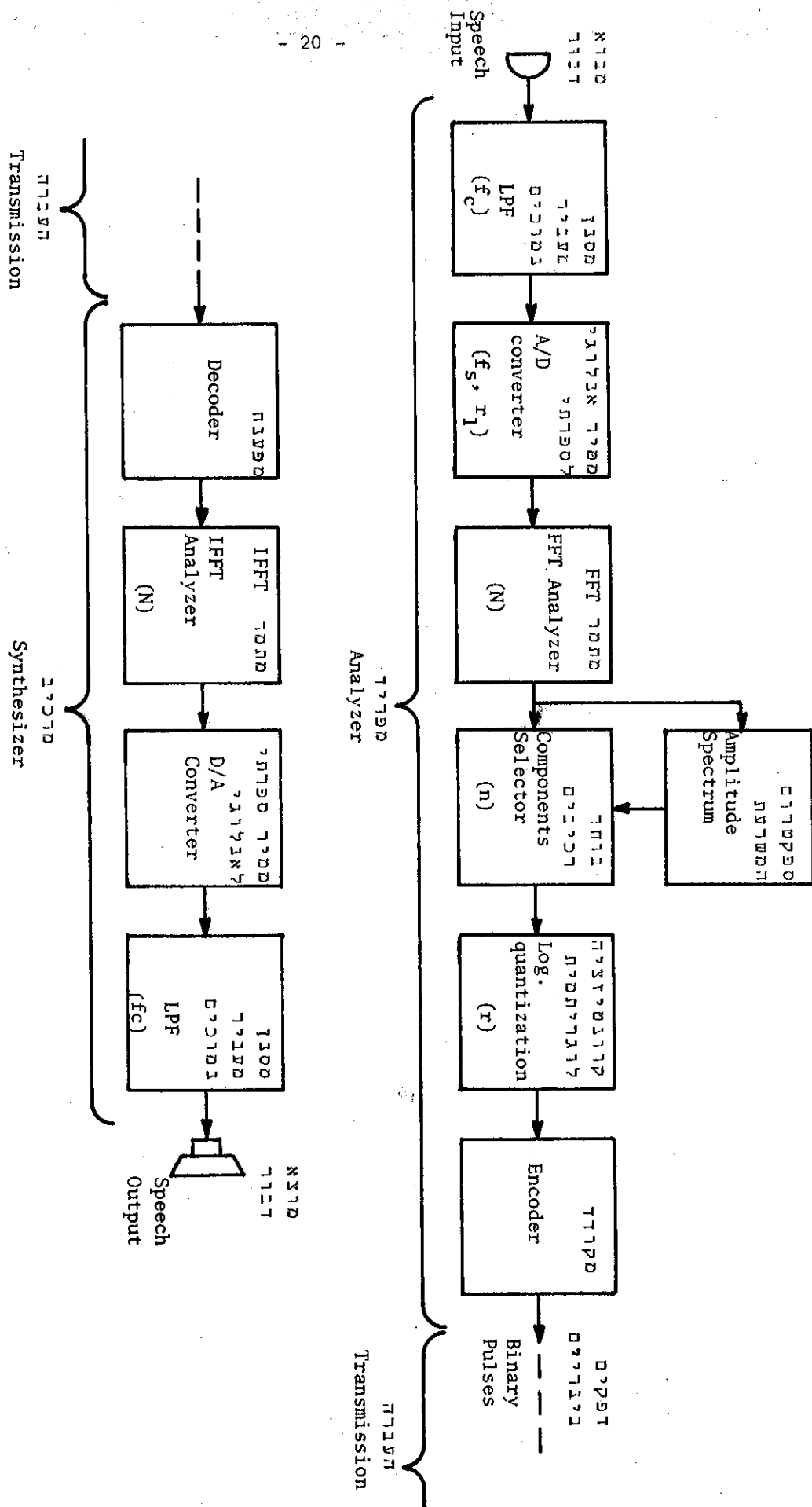


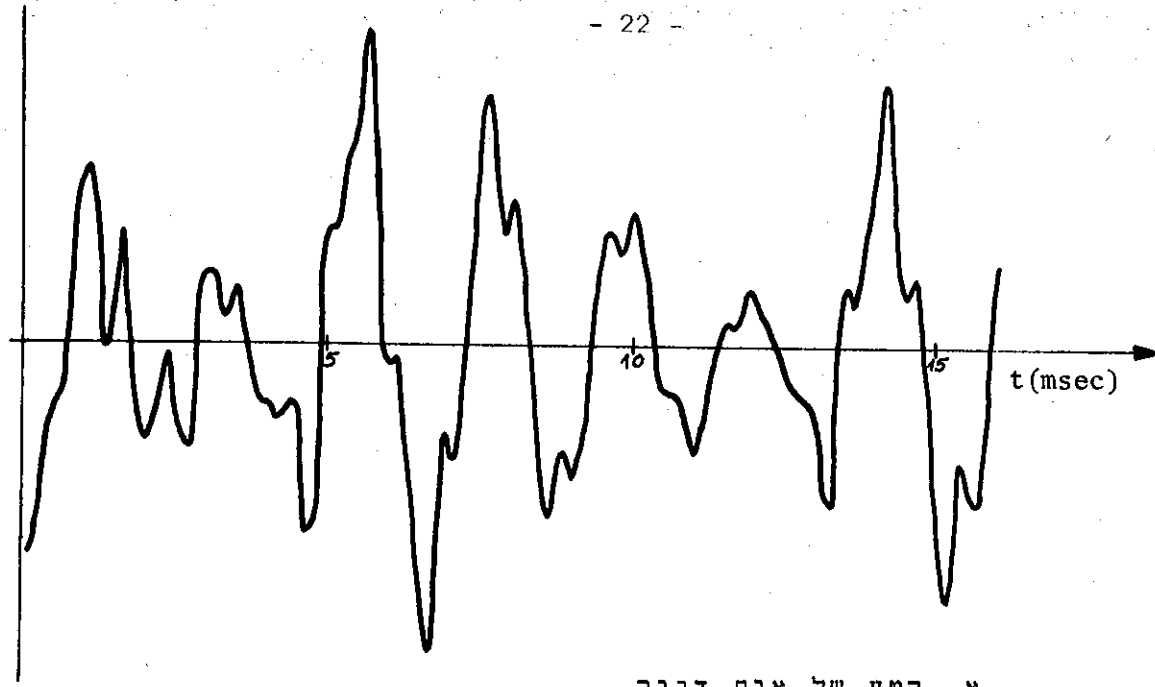
Fig. 4 Vocoder based on local maxima method

סדרת הדגימות מועברת למישור החדר באמצעות מחמר FFT הפועל על בלוקים של N דגימות. תפוקתו היא N רכיבים קומפלכסים של החמרת פוריה הדיסקרטית. אך כפי שהוזכר קודם רק $N/2$ מן הרכיבים הם בעלי ענין עבורנו היות ו- $N/2$ האחרים נושאים אותו מידע. בכדי לבצע בחירה של הרכיבים הדומיננטיים מחשבים תחילה את ספקטרום המשרעת בכל בלוק. מתוך ספקטרום זה בוחרים את כל הקוים המהוים שיאים מקומיים (local maxima) של פונקציית הספקטרום ומחוכם בוחרים הקוים הגבוהים ביותר. בוחר הרכיבים מעביר הלאה רק את אותם רכיבי החמרת פוריה המתאימים ל- n הקוים הנבחרים. הנסיון מוכיח כי חאור את הדבור ע"י רכיבי החמרת פוריה המתאימים לשיאים בפונקציית הספקטרום מבטיח תאור טוב של מעטפת הספקטרום.

הרכיבים הנבחרים עוברים קוונטיזציה לוגריתמית של משרעתם לפי x סביות. הקוונטיזציה הינה לוגריתמית כדי להבטיח חלוקה עדינה של הרמות גם עבור הרכיבים שעצמתם קטנה, ביחס לעצמתם של מספר רכיבים בולטים (של הפורמנט הראשון בדרך כלל), אך תרומתם לאיכות הדבור ניכרת. בקוונטיזציה לוגריתמית יהיו הבדלים קטנים בין רמות קוונטיזציה עוקבות עובר עצמות אות נמוכות והבדלים גדולים יותר, בין רמות עוקבות, ככל שעצמת האות הולכת וגדלה. עדיפות הקוונטיזציה הלוגריתמית יש לה גם ביטוס תיאורטי היות ובממוצע פונקציית צפיפות ההסתברות (probability density function) של ספקטרום הדבור קרובה יותר ללוגריתמית מאשר ללנארית. ידוע, ששגיאה רבועית ממוצעת מינימלית, תתקבל כאשר רמות הקוונטיזציה יהיו במרווחים כאלה, שכל רמה תתרום תוספת הסתברות זהה, לפלוג ההסתברות של הדגימה. על בן חלוקה לוגריתמית של הרמות תקטין את השגיאה [17, 13]. היות ולא כל רכיבי ההחמרה מועברים יש לתאר את מקומו במשור החדר (מספרו הסדורי) של כל רכיב נבחר ולהעביר נתון זה למרכיב. פעולה זו מבוצעת במקודד, אשר מעביר אל המרכיב סדרה של דפקים בינריים המתארים לגבי כל רכיב את מקומו, את המשרעת של החלק המדומה והמשרעת של החלק הממשי. ההעברה בין המפריד למרכיב היא ספרתית.

המרכיב כולל מפענח, מחמר IFFT, ממיר ספרתי לאנלוגי (D/A converter) ומסנן מעביר נמוכים. האות הספרתי המגיע מן המפריד עובר פענוח כך שבמבוא ממיר IFFT יהיו הרכיבים הנבחרים בעלי משרעת כפי שנמסרה מן המפריד ואילו שאר הרכיבים יאופסו. ממיר IFFT פועל גם הוא על בלוקים של N דגימות. תפוקתו של ממיר זה תהיה הדגימות הספרתיות של אות הדבור המשוחזר. כדי לקבל מהן אות דבור אנלוגי נמסרת תפוקת הממיר לממיר ספרתי לאנלוגי ומשם לאחר מסנן מעביר נמוכים בעל תדר קטעון f_c , מתקבל הדבור המשוחזר.

להדגמת שיטת הבחירה של הרכיבים הדומיננטיים, בהחמרת פוריה, נחבונן בציור מס' 5. בציור מס' 5 מתואר ספקטרום צר-זמן של קטע הדבור באורך 16 מילישניות אשר הצורה הזמנית שלו מופיעה בציור מס' 5א. אות הדבור נדגם בתדר של 8 ק"ה, והבלוק כולל 128 דגימות, לכן תחום התדרים 0 עד 4 ק"ה מתואר ע"י 64 קוים ספקטרליים, עם רזולוציה של 62.5 הרץ. היות ואות הדבור היה מוגבל בתחום התדרים 300 הרץ עד 3 ק"ה הרי הקוים 0 עד 4 (מתאים ל-300 הרץ)

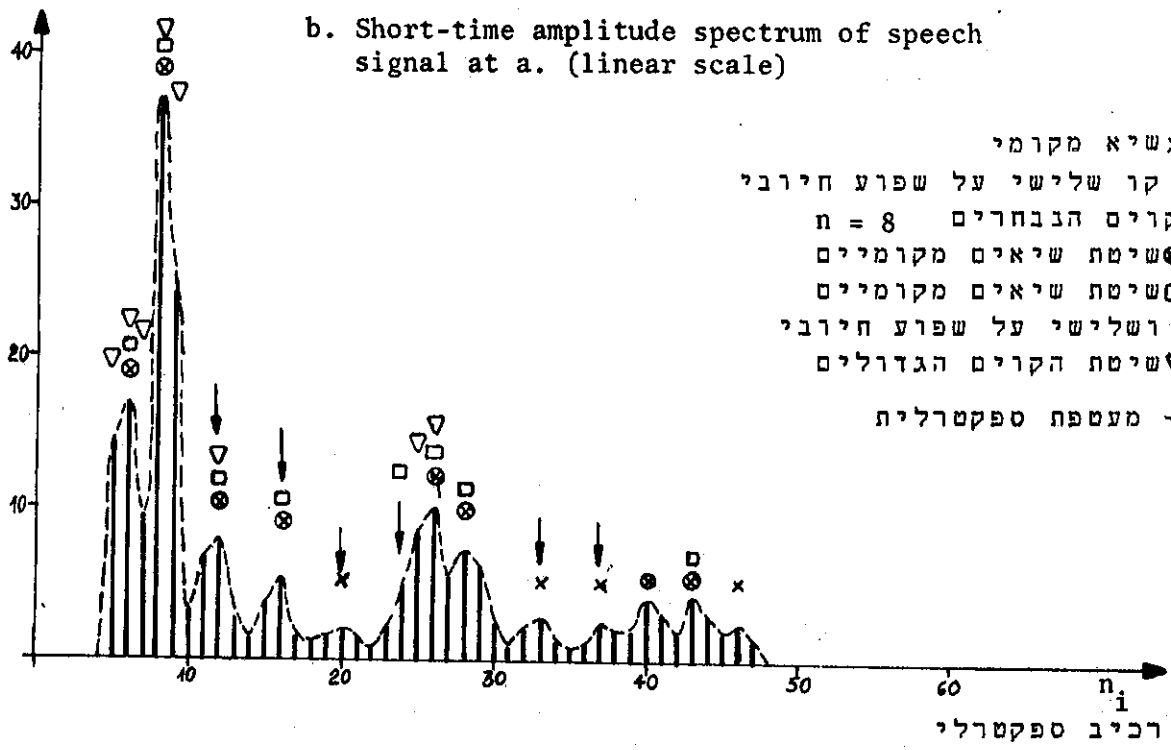


א. קטע של ארת דבור

a. Speech signal

ב. ספקטרום משרעת צר-זמן של ארת הדבור בא. (סקלה לבארית)

b. Short-time amplitude spectrum of speech signal at a. (linear scale)



א שיא מקומי
 ↓ קר שלישי על שפוע חירובי
 הקוים הנבחרים $n = 8$
 ● שיטת שיאים מקומיים
 □ שיטת שיאים מקומיים
 ושלשי על שפוע חירובי
 ▽ שיטת הקוים הגדולים

--- מעטפת ספקטרלית

ציור מס. 5 הדגמה של שיטות בחירה שונות
 Fig. 5 Examples for different selecting methods

ו-48 (מתאים ל-3 ק"ה) עד 64 (מתאים ל-4 ק"ה) הנם אפס. הקוים המסומנים ב-X הנם הקוים המהוים שיאים מקומיים של פונקצית הספקטרום (קוים מס' 6, 8, 12, 16, 20, 26, 28, 33, 37, 40, 43, 46). מתוך הקוים האלה נבחרו שמונה הקוים הגדולים ביותר, הם סומנו ב-X (קוים מס' 6, 8, 12, 16, 26, 28, 40, 43). רכיבי החמרה פוריה הדיסקרטית המתאימים לקוים אלה יבחרו לייצג את ההחמרה של קטע זה ואילו שאר הרכיבים יאופסו, כפי שרואים מן הדוגמה הרי הקוים שנבחרו מתארים בנאמנות רבה את הספקטרום ולכל פורמנט יש ייצוג בין הקוים הנבחרים. (הסמונים הנוספים המופיעים בציור מס' 5ב' מחיתקים לשיטות בחירה שונות אשר יוסברו בהמשך החבור).

איכות הדבור המשוחזר וקצב ההעברה חלויים במספר הדגימות בבלוק-N, מספר הקוים הנבחרים-n, ומספר הסביות בקוונטיזציה-x. להבטחת קאור נאמן של הספקטרום (ואיתו איכות גבוהה של הדבור) רצוי שהרזולוציה תהיה גבוהה (Δf קטן), מה שמחייב אורך בלוק גדול לפי (12). אך אורך הבלוק מוגבל ע"י מידת הסטציונריות של פונקצית הערוור וכן רצוי שלא יעלה על משכן של הפונמות או ההברות. על כן יהיה אורך הבלוק מוגבל לכמה עשרות מילישניות (סעיף 4.1.2). כדי להגדיל את הצמצום המושג בשיטה זו, כלומר להקטין את קצב ההעברה, רצוי להקטין את n ואת x. ברור עם זאת שאיכות הדבור המשוחזר יחסית הפוך ל-n ול-x (החלות אינה לנארית). גדלם של n ו-x יקבע כפשרה בין דרישות מנוגדות אלה.

בשיטה המוצעת ניחן יהיה להשיג קצבי העברה הנמוכים מ-6000 סביות לשניה עם איכות טובה של הדבור המשוחזר.

5.3.2 שיטת הקוים הגדולים

Shaffer, Manley, Luck [16] הציעו שיטה לצמצום תכולת אות הדבור, המבוססת גם היא על צמצום במשור החדר, אליו עוברים באמצעות FFT. מתוך N הרכיבים של החמרה בוחרים n הרכיבים הגדולים ביותר ושאר הרכיבים מאופסים. האמפליטודה והמופע של הרכיבים הנבחרים מועברת למרכיב ושם משוחזר אות הדבור ע"י החמרה הפוכה (IFFT). הדבור המשוחזר דומה באיכותו לדבור המקורי כאשר שומרים על רבע מרכיבי החדר. אורך הבלוק בנסויים אלה היה 256 דגימות.

להדגמה שיטת בחירה זו נתבונן שוב בציור מס' 5ב' נניח שברצוננו לבחור 8 הקוים הגדולים ביותר. הקוים האלה סומנו במשולש ∇ (קוים מס' 5, 6, 7, 8, 9, 12, 25, 26). רואים בברור שכשיטה זו מרכיב הקוים שנבחרו הם סמוכים זה לזה ומתארים את הפורמנט הראשון בלבד. מעוטם של הקוים מייצג את הפורמנט השני ואף לא קו אחד עבור הפורמנט השלישי. לעומת זאת בשיטת השיאים המקומיים נבחרו קוים לאורך כל חחום הספקטרום

והם מחארים את כל שלושת הפורמנטים. תאור הספקטרום המצומצם בשיטת השיאים המקומיים נראה נאמן יותר למקור. מסיבה זו נראה שבשיטת השיאים המקומיים ניתן להשיג איכות דבור טובה עם צמצום גדול יותר, המגיע עד $1/32$ מרכיבי הספקטרום ולא רק רבע כפי שזה נמסר ביחס לשיטת הקוים הגדולים. גם השמוש בקוונטיזציה לוגריתמית ולא לנארית כמו אצל Luck מבטיח תאור נאמן יותר של הרכיבים אשר משרעחם אמנם נמוכה אך תרומתם לאיכות הדבור ניכרת.

5.4 צמצום אותות דבור באמצעות התמרות אורתוגונליות שונות

מלבד התמרת פוריה שמשו גם התמרות אורתוגונליות אחרות לנסיונות של צמצום אותות דבור באמצעות עיבוד ספרתי. התמרת Walsh-Hadamard מושכת במיוחד בגלל הנוחיות לממושה בחמרה, הנובעת מכך שאין צורך בכפל. Robinson ו-Campanella [3] ערכו השוואה של שלוש התמרות, אורתוגונליות דסקרטיות, כאמצעי לצמצום אותות דבור; התמרת פוריה, התמרת Walsh-Hadamard והתמרת Karhunen-Loeve השיטה המוצעת על ידם מבוססת על כך שהשונות (variances) של רכיבי ההתמרות האורתוגונליות שונים עבור רכיבים שונים. כללית השונות קטנה ככל שמספר הרכיב עולה. על כן מספר רמות קוונטיזציה שיש להקציב לכל רכיב של ההתמרה יורד ככל שיורדת השונות. בשיטה המתוארת הוקצב לכל רכיב של ההתמרה מספר סביות לקוונטיזציה, בהתאם לשונות שלו, כאשר הקריטריון היה שמירה על יחס אות לרעש קוונטיזציה הזחה לזה המחקבל באפנון דופק מקודד (PCM). הצמצום המושג בשיטה זו איננו גדול (7.5 k bits/sec עד 13.5 kbits/sec מחוך 56 kbits/sec), אך ניתן ללמוד מכאן שההתמרה היעילה ביותר היא התמרת Karhunen-Loeve ואילו התמרת Walsh-Hadamard הנה הפחות יעילה מכולן.

Elliott, Shum ו-Brown [23] הציעו שיטה לצמצום המבוססת על התמרת Walsh-Hadamard ובעיקרה דומה לשיטת הקוים הגדולים בהתמרת פוריה (סעיף 5.3.2). מבצעים התמרת Walsh-Hadamard על בלוקים של 64 דגימות של אות הדבור ובוחרים שמונה או ארבע רכיבים הגדולים ביותר. שחזור הדבור מתבצע ע"י התמרת Walsh-Hadamard הפוכה של הרכיבים הדומיננטיים. קצבי ההעברה הנחוצים בשיטה זו הם 6.5 kbits/sec עד 13 kbits/sec ומתקבל דבור באיכות מספקת.

6. תאור העבודה ותוצאותיה

6.1 הדמוי במחשב

המחקר בוצע ע"י דמוי, של השיטה המוצעת לצמצום ספרתי של אותות דבור, במחשב ספרתי IBM 360/44 אשר כלל יחידת ממיר אנלוגי לספרתי וספרתי לאנלוגי כמסוף למחשב. מערכת המחשב כללה גם קורא כרטיסים מנוקבים, מדפסת, זכרון דיסקה מגנטית וזכרון סרט מגנטי. התכניות נכתבו בפורטרן IV (Fortran IV). התכניות שנוצלו בעת הדמוי היו:

- א. חכנית דגימה והמרה אנלוגית לספרתית.
- ב. חכנית ראשית לבצוע העבוד המוצע.
- ג. חכנית העברה מסרט מגנטי לדיסק מגנטי.
- ד. חכנית המרה ספרתית לאנלוגית.

הדפסי התכניות מובאים בנספח ד'.

לצורך המחקר הוקלטו 5 סדרות של מלים עבריות קצרות בנות שתי הברות. כל סדרה כללה 15 מלים (פרט לראשונה שכללה 11) ונקרא לה להבא לוח. פירוט המלים בלוחות השונים מופיע בנספח ב'. היה צורך במספר לוחות היות ובנסווי השמעה לצורכי הערכת איכות הדבור אפשר היה להשמיע כל לוח רק פעם אחת לכל מאזין בשל למידת הלוח על פה ע"י המאזין. כמו כן שמוש במספר לוחות אפשר השואה בין טקסטים שונים. הלוחות הוקראו והוקלטו על קלט סרט Revox במהירות של $7\frac{1}{2}$ אינץ' לשניה (IPS).

אורך הטקסט הוגבל ל-17 שניות בגלל זמן המחשב המוגבל שעמד לרשות עבודה זו. עבוד קטע של 17 שניות (מהכנסת אוח אנלוגי ועד הוצאת אוח אנלוגי מעובד) נמשך כ-10 דקות.

היות ולא היה ידוע למחבר (בזמן עריכת הנסויים) על לוחות מאוזנים פונטית בעברית או לוחות כלשהם אחרים בעברית המקובלים לעריכת נסויי מובנות, חוברו במסגרת עבודה זו לוחות של מלים בנות שתי הברות, תוך נסיון לכלל את מרבית הצלילים של השפה. הדמוי במחשב נערך בדרך הבאה:

- א. המרת אוח הדבור האנלוגי לספרתי ואכסון של סרט מגנטי. המעריך לקלט אנלוגי למחשב מחואר בציור מס' 6. אוח הדבור מקלט-סרט הועבר דרך מסנן מעביר תחום, למגבר אשר עצמתו כוונה כך שמתח המוצא שלו לא יעלה על $\pm 5V$ היות ואלה הגבולות של רמת הכניסה לממיר אנלוגי לספרתי. הפקוח על כך בוצע באמצעות התצוגה. הפוקח המגבר נמסרה לממיר אנלוגי לספרתי לדגימה והמרה, אך הממיר היה מוגבל לקצב דגימה של 4 ק"ה בלבד. על כן היה צורך להשמיע את הדבור מקלט-סרט במחצית ממהירות ההקלטה וכך להקטין פי שתיים את תחום התדרים של אוח הדבור. תחום התדרים של אותות הדבור בכל הנסויים הוגבל ל-300 הרץ עד 3 ק"ה שהוא תחום התדרים המקובל במערכות תקשורת דבור. עקב

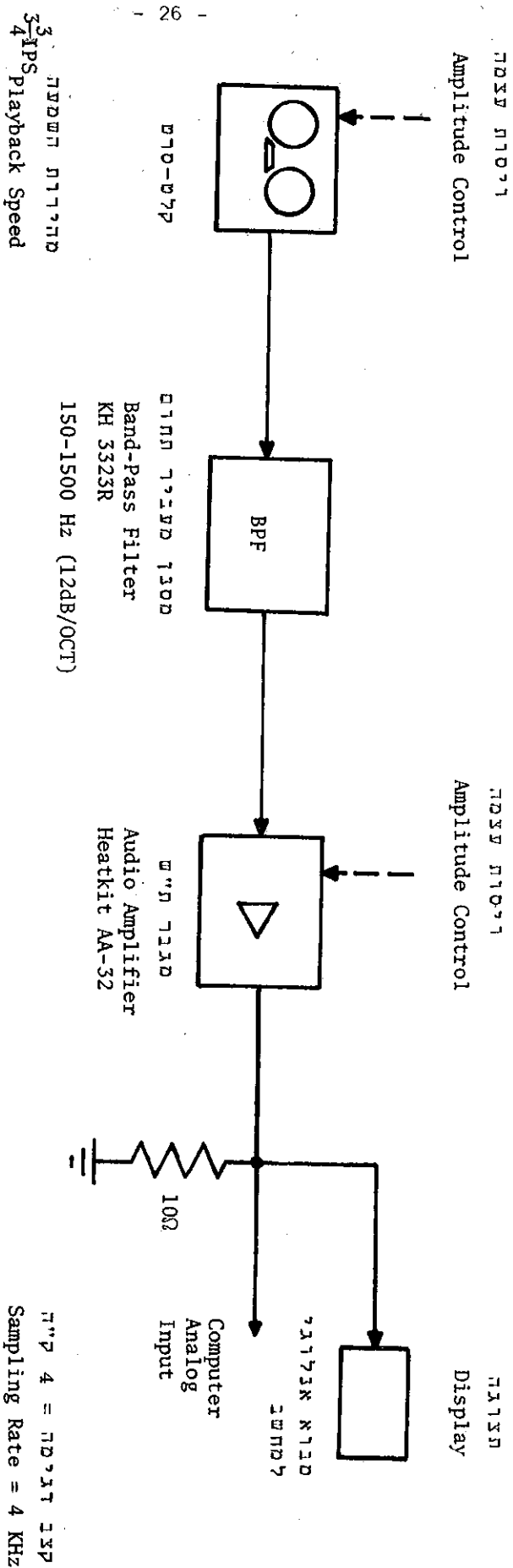


Fig. 6 Set-up for analog input to computer

השמעה במחצית המהירות כוון המסנן לתחום תדרים 150 הרץ עד 1.5 ק"ה. הקוונטיזציה בממיר היחה ע"י 15 סביות בחוספת סיבית סימן. לאחר שהאות האנלוגי הומר לספרתי הוא אוכסן תחילה על דיסקה מגנטית ומשם הועבר לסרט מגנטי לשמירה והמשך העבוד.

ב. בשלב השני הוכנסה תכנית לעבוד ספרתי של האותות המאוכסנים. העבוד היה לפי השיטה שהוצעה בסעיף 5.3.1, כאשר מדי פעם משנים את הפרמטרים השונים לבדיקת השפעתם. תוך כדי עבוד האותות הודפסו במדפסה המחשב רכיבי החמרת פוריה והספקטרום לפני ואחרי העבוד וכן הודפסו בצורת גרפים האותות המקוריים והמעובדים במשור החדר ובמשור הזמן. הדפסה זו בוצעה ביחס למספר מצומצם של בלוקים נבחרים והיא איפשרה פקוח יעיל על העבוד וכן ספקה נתונים על תהליך העבוד. בתום העבוד אוכסנו האותות על סרט מגנטי שני.

ג. בשלב האחרון הועברו האותות המעובדים מסרט מגנטי לדיסקה מגנטית ומשם דרך ממיר ספרתי לאנלוגי החוצה לקלט-סרט אנלוגי. המערך לפלט אנלוגי ממחשב מתואר בציור מס' 7. האות האנלוגי במוצא הממיר ספרתי לאנלוגי הועבר דרך מסנן מעביר תחום ומשם לקלט סרט אנלוגי לצורך הקלטה. היות והממיר ספרתי לאנלוגי עובד בקצב 8 ק"ה הרי הפלט נעשה בזמן אמיתי והוקלט על כן במהירות הרגילה של קלט-הסרט (7.5 IPS). המסנן העביר את תחום התדרים 300 הרץ עד 3 ק"ה וסילק רעשים ואותות בלתי רצויים שנחווספו. התצוגה במוצא המחשב היחה לצרכי פקוח ואתור תקלות בלבד.

האות המעובד שהוקלט, שימש לבצוע נסויים להערכת איכות הדבור המשוחזר, בשיטה המוצעת. ס"ה נערכו 26 הרצות במחשב אשר כללו 30 נסויים שונים. פרטי הנסויים השונים סוכמו בטבלה A. 1 אשר בנספח א'. הנסויים מוספרו ונתיחס אליהם להנא לפי מספור זה. בנסויים אלה היו N, x, y, z , שיטת הבחירה וצורת "חלון הזמן" פרמטרים שהשפעתם על מובנות הדבור נבדקה.

6.2 בדיקות מובנות

המובנות שמשו כקריטריון בסיסי להערכת השיטה המוצעת לצמצום אות הדבור ולהערכת איכות הדבור המשוחזר. בעזרת קריטריון המובנות נבדקה השפעת פרמטרים שונים. המובנות נמדדה בבדיקות אשר נערכו על טקסטים מוקלטים, אשר היו קטעי דבור משוחזר (לאחר צמצום), כפי שנחקבל בנסויים השונים.

נחאר להלן את השיטה בה נערכו בדיקות המובנות. הטקסט המוקלט הושמע מתוך הקלט-סרט דרך אזניות. המאזין נתבקש לרשום כל מלה או כל חלק מילה שהוא שומע. ההשמעה נערכה בקצב איטי יחסית, כדי לאפשר למאזין להספיק לרשום, כל מלה שהוא שומע, לפני שמושמעת

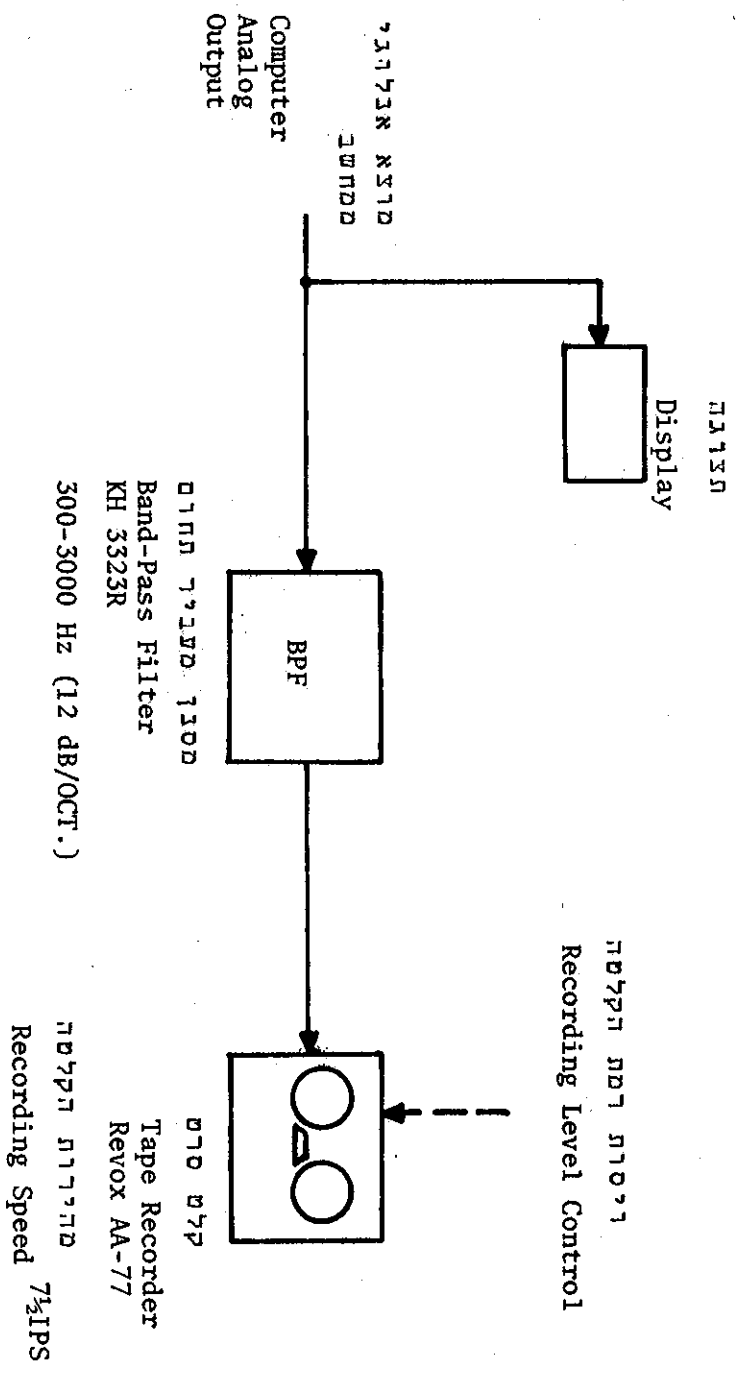


Fig. 7 Set up for analog output from computer

מילה חדשה. לוח מסוים הושמע לכל מאזין רק פעם אחת וזאת כדי להמנע מקבלת תוצאות המושפעות משמיעה קודמת. היות ומספר המאזינים שעמד לרשות נסויים אלה היה מוגבל היה צורך להחליף מדי פעם את הקטע אשר שימש לעיבוד. חלק מן המאזינים היו מאזינים "תמימים" (ללא כל נסיון קודם בנושא זה) וחלקם היו מאזינים מנוסים. כדי להכין כל מאזין לצורת הבדיקה ולנדרש ממנו הוסבר לו, לפני תחילת הבדיקות שהוא ישמע קטע דבור הכולל מלים עבריות קצרות אשר עברו עבוד מסוים. עליו להשתדל לרשום את אשר הוא שומע. כמוכן תורגל כל מאזין ע"י השמעת קטע דומה ורשום הנקלט. המאזינים לא ידעו מה הן המלים שישמעו ומתוך איזה מבחר של מלים הוצאו, כלומר מבחינתם גודל המילון היה בלתי מוגבל. רק לאחר מכן הושמע הטקסט שעמד למבחן.

המלים כפי שנרשמו ע"י המאזין הושו, על בסיס של הברות, לטקסט המקורי. המובנות עבור כל מאזין נקבעה כאחוז ההברות שנרשמו נכונה מתוך ס"ה ההברות אשר הושמעו. המובנות נמדדה על בסיס של הברות היות ומדידה זו רגישה מאד לפגיעה באיכות הדבור, בנגוד למדידה בעזרת משפטים למשל. קביעה זו נכונה במיוחד לתחום בו אין פגיעה רצינית במובנות היות ואז יקלטו המשפטים בשלמותם, מתוך ניחוש והשלמה מתוך הקשר, וזאת למרות חוסר קליטה של חלק מן ההברות (French [10]). ההברות בלוחות שחברו היו חלק ממילים ולא הברות סתמיות, היות והברות סתמיות אינן נושאות כל מובן ולכן אין הן יכולות לשמש למדידה מובנות (Lehiste [15]).

המובנות בכל ניסוי נקבעה כממוצע תוצאות המובנות אשר הושגו ע"י המאזינים השונים באותו ניסוי. בכל ניסוי השתתפו לפחות 10 מאזינים. התוצאות של בדיקות המובנות עבור המאזינים השונים מרוכזות לפי הנסויים בטבלה A.2 אשר בנספח א'. המאזינים סומנו במספרים שוטפים וכך הם מופיעים בטבלה זו. באותה טבלה מופיעה גם המובנות הממוצעת עבור כל ניסוי. תוצאות אלה משמשות לכל הניתוחים וההערכות להלן. סך הכל נעזרתי ב-105 מאזינים ונערכו 328 בדיקות מובנות.

המטרה למצא את פרמטרי שיטת הצמצום המבטיחים צמצום מכסימלי עם מובנות הברות של כ-70 אחוז (שהיא שוות ערך למובנות משפטים של 90 עד 95 אחוז).

6.3 תלות המובנות בשלושה פרמטרים בסיסיים

שלושת הפרמטרים הבסיסיים בהם תלויה המובנות בשיטה המוצעת הם: מספר הרכיבים הנבחרים-N, מספר הסביות בקוונטיזציה של משועת הרכיבים-x ואורך הבלוק לעבוד-N. נחאר להלן תלות המובנות בכל פעם בפרמטר אחד כאשר שאר שני הפרמטרים קבועים וכמובן קבוע הלוח כדי שהחשואה תהיה בלתי תלויה בטקסט. הבחירה של הרכיבים היתה בשיטה שיאים מקומיים.

6.3.1 תלות המובנות במספר הרכיבים הנבחרים

בטבלה מס' 1 להלן מופיעות תוצאות נסויים שונים אשר נערכו עם לוח מס' 1 ובהם שונה מספר הרכיבים הנבחרים אך אורך הבלוק ומספר הסביות בקוונטיזציה נשמרו קבועים:

$$N = 512 \quad r = 6$$

טבלה 1: תלות המובנות במספר הקוים הנבחרים

Table 1. Intelligibility vs. no. of chosen lines

מספר רכיבים נבחרים No. of chosen lines	n	256	64	32	16	12	8	4
מובנות Intelligibility	I(%)	96	85.3	79	73.9	65.2	54.2	14.6

בציור מס' 8 מתוארת תלות זו בגרף, גרף זה מצטיין בשלושה אזורים בולטים:

א. אזור רייה בו המובנות גבוהה מ-85 אחוז ומספר הרכיבים הנבחרים גדול מ-64.

ב. אזור הברך בו המובנות בין 65 אחוז ל-85 אחוז ומספר הרכיבים הנבחרים בין 12 ל-32.

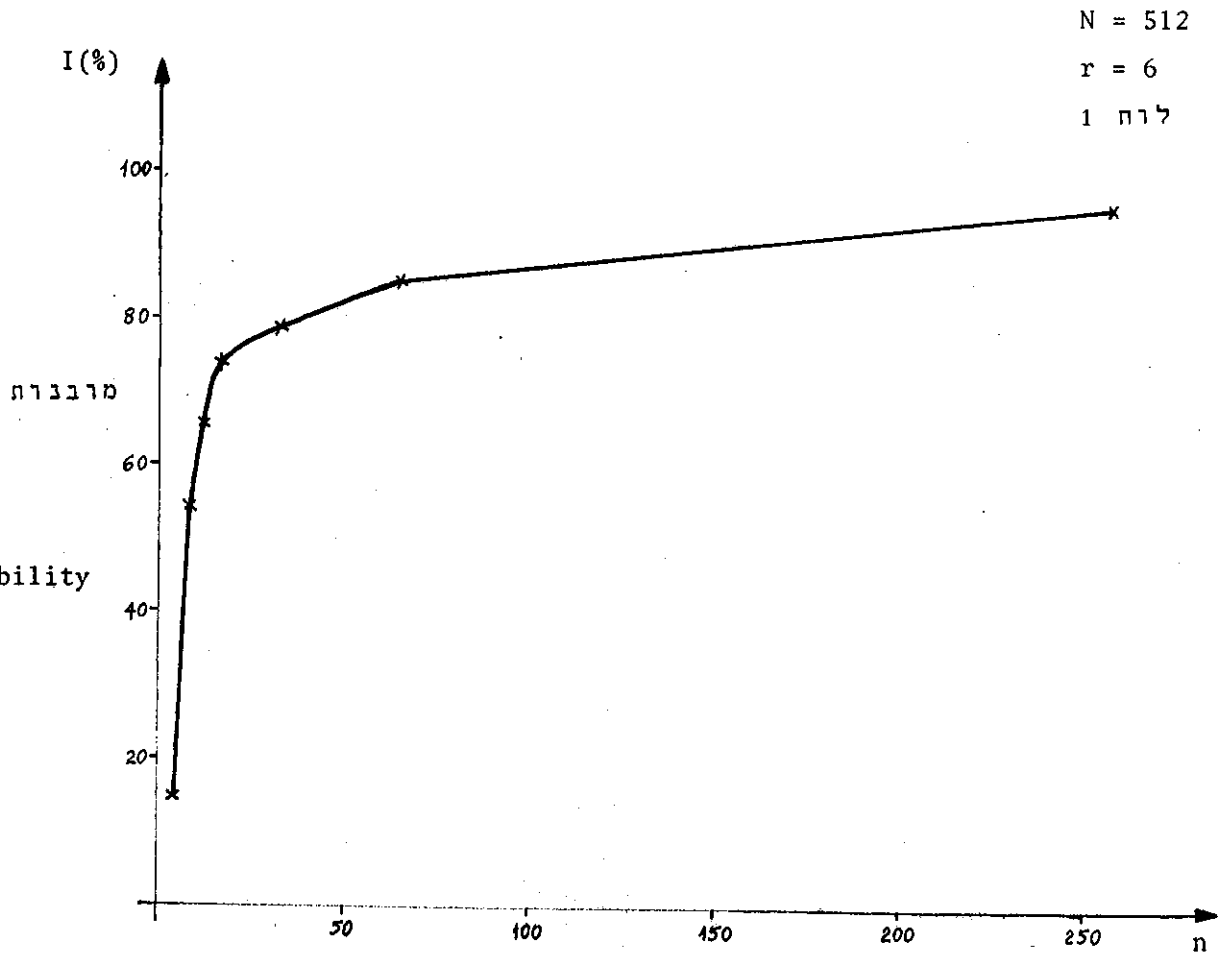
ג. אזור שלישי בו המובנות נמוכה ונופלת בצורה חריפה עם ירידה מספר הרכיבים.

מתוך בחינה של הגרף ניתן לאמר שהנקודה האופטימלית היא נקודת השבירה של הגרף והיא עבור 16 = n (כלומר נותרו 1/32 רכיבים מס' רכיבי החמרה). בנקודה זו מקבלים גם מובנות טובה של 74 אחוז. כל נסיון לצמצם מספר הרכיבים מעבר ל-16 גורם לירידה ניכרת במובנות. כל נסיון לשפר המובנות בצורה ניכרת מצריך העלאה גדולה במספר הרכיבים הנבחרים.

6.3.2 תלות המובנות במספר רמות קוונטיזציה

בטבלה מס' 2 להלן מופיעות תוצאות נסויים שונים אשר נערכו עם לוח מס' 3 ובהם שונה מספר הסביות בקוונטיזציה של משרעת הרכיבים הנבחרים, אך נשמרו קבועים אורך הבלוק ומספר הקוים הנבחרים:

$$N = 512 \quad n = 16$$



מספר הקווים הספקטראליים הנבחרים
No. of the chosen spectral lines

ציור מס. 8 תלות המובנות במספר הקווים הנבחרים

Fig. 8 Intelligibility as a function of the no. of chosen lines

טבלה 2: תלות המובנות במספר רמות קוונטיזציה

Table 2. Intelligibility vs no. of quantization levels,

מספר סביות קוונטיזציה No. of quantization bits r	2	3	4	5	6	8
מובנות Intelligibility $I(\%)$	27.0	58.7	66.9	67.1	67.4	77.5

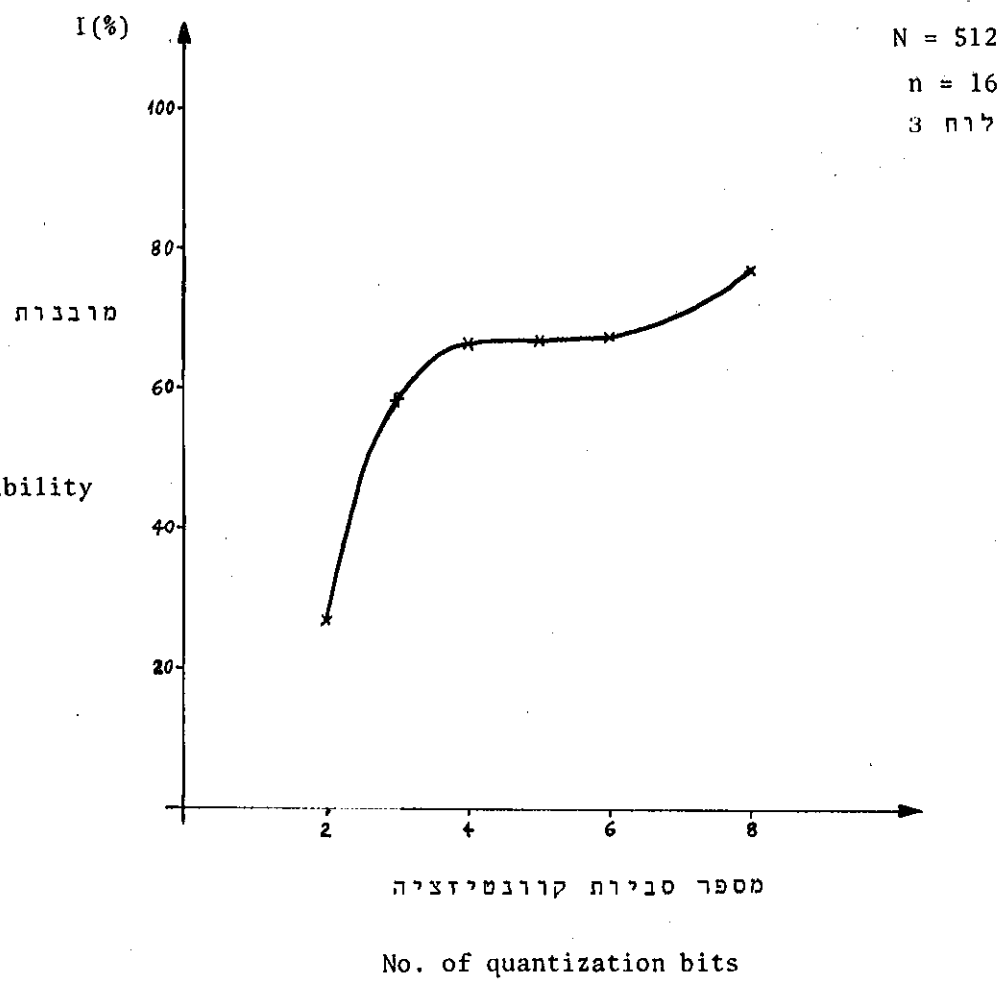
נבחר שהמספר r מציין מספר הסביות בקוונטיזציה של הערך המוחלט של משרעת כל חלק של הרכיב. ובנוסף לו קימת סבית אחת של סימן כלומר ס"ה $r+1$ סביות נחוצות לתאור המשרעת של כל חלק של רכיב. בציר מס' 9 מחוורת תלות זו בגרף. מתוך התבוננות בגרף זה נראה שאין הבדל ניכר בין שמוש בארבע סביות לבין שמוש בשש סביות היות ובשני המקרים המובנות היא כ-67 אחוז, לכן כדאי לבצע קוונטיזציה עם ארבע סביות בלבד. במדה ונרצה להגדיל המובנות יש לעבור ל-8 סביות. הקטנת מספר הסביות מתחת לארבע לא נראה מעשי היות והמובנות נפגעת מאד.

6.3.3 תלות המובנות באורך הבלוק

בטבלה מס' 3 להלן מופיעות תוצאות נסויים שונים אשר נערכו עם לוח מס' 3 ובהם שונה אורך הבלוק ע"י שנוי מספר הדגימות לבלוק אך נשמר קבוע מספר רמות הקוונטיזציה וכן נשמר היחס בין מספר הרכיבים הנבחרים למספר הדגימות לבלוק (יחס זה נשמר ליתר דיוק כך שקצב ההעברה ישאר קבוע):

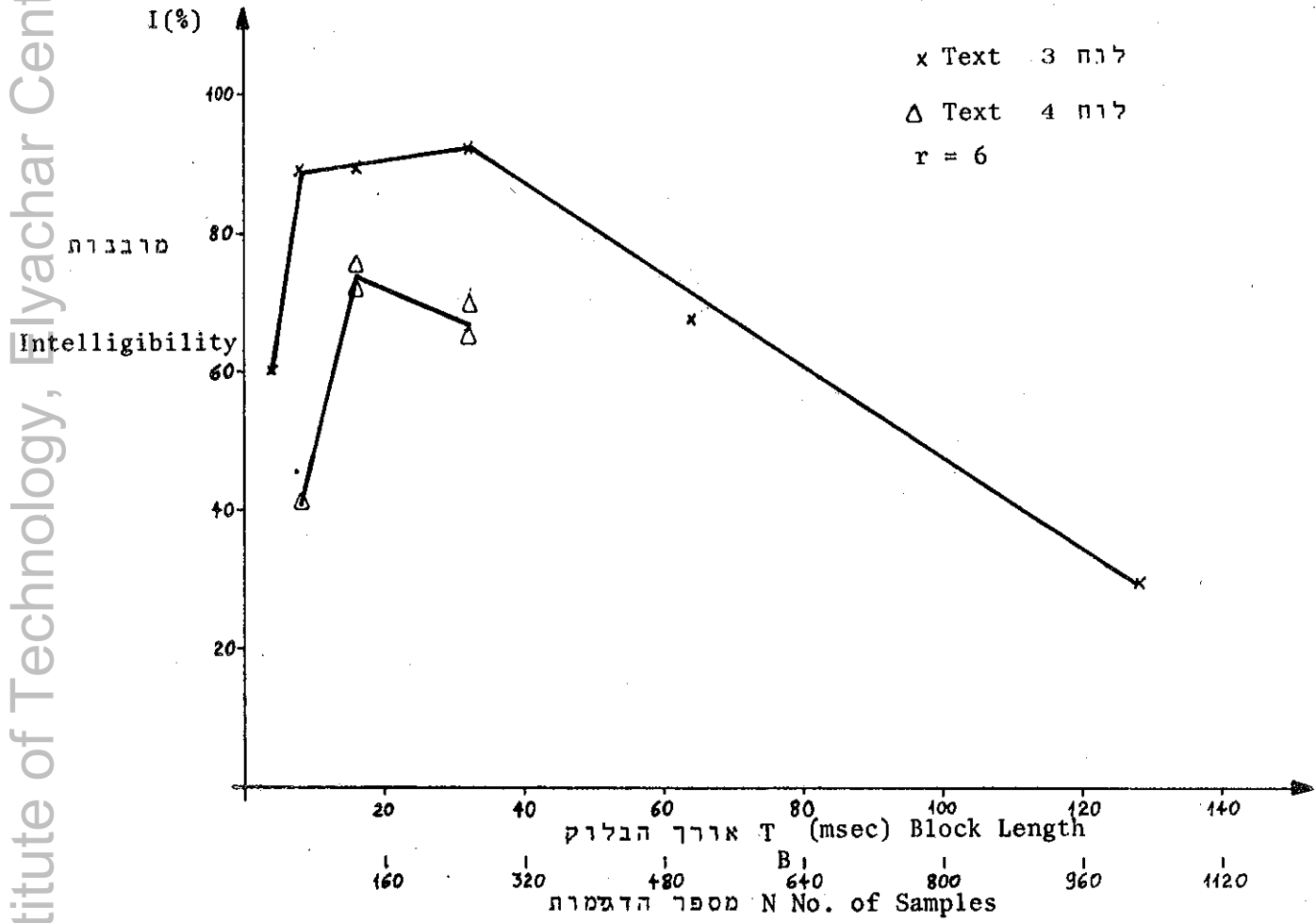
$$r = 6 \quad n/N = 1/32$$

תלות המובנות באורך הבלוק מתוארת בציר מס' 10. היות ונתקבלה מובנות גבוהה בשלושה ארכי בלוק שונים (8 עד 32 מילישניות), נערכו נסויים נוספים באותם חנאים, עם לוח מס' 4. תוצאות נסויים אלה מופיעות גם הן בטבלה מס' 3 ובציר מס' 10. בשני ארכי בלוק בהם נתקבלו תוצאות גבוהות גם בלוח זה נבדקה המובנות גם כאשר מקטינים את n באחד ומתקבל המצב בו נשארים בדיוק $1/32$ רכיבים. ארכי בלוק אלה הם 16 ו-32 מילישניות ונוסו 4 ו-8 קוים בהחאמה. מתוך הגרפים מתברר שארכי הבלוק המועדפים הם 16 ו-32 מילישניות, הכוללים 128 ו-256 דגימות לבלוק בהחאמה.



ציר מס. 9 תלות המרבנות במספר רמות קוונטיזציה

Fig. 9 Intelligibility as a function of the no. of quantization level



ציור מס. 10. תלות המובנות באורך הבלוק

Fig.10 Intelligibility as a function of Block length

טבלה 3: תלות המובנות באורך הבלוק

Table 3. Intelligibility vs. Block length

אורך הבלוק Block length T_B (msec)	מספר הדגימות No. of samples N	מספר קוים נבחרים No. of chosen lines n	Intelligibility I(%)	
			לוח 3 Text 3	לוח 4 Text 4
128	1024	31	29.4	-
64	512	16	67.4	-
32	256	9	92.1	69.3
32	256	8	-	64.7
16	128	5	89.4	75.0
16	128	4	-	71.7
8	64	2	89.2	41.1
4	32	1	60.2	-

מחוך התוצאות הנסויים עם לוח 4 נראה אורך בלוק של 16 מילישניות כטוב ביותר היות והמובנות בתנאי זה עולה בכ-6 אחוז על המובנות בבלוק של 32 מילישניות. מסקנה זו סותרת אמנם את אשר התקבל בנסויים עם לוח 3 אך חשוב לשים לב שבמקרה זה ההבדלים היו קטנים מאד וזאת ברמת מובנות גבוהה מאד. לכן נראה שהחלטה לפי לוח 4 תהיה נכונה יותר. ההבדל ברמות המובנות בין שני הלוחות (באותם תנאים) נובע כנראה מכך שלוח 4 קשה יותר לקליטה היות והוא כולל מלים הנפוצות פחות מאלה בלוח 3.

עובדה מעניינת היא שאורך הבלוק האופטימלי הוא בסדר גודל כזה שמבטיח סטציונריות של פונקצית הערור (פחות מ-40 מילישניות) וכמו כן הוא בסדר גודל של זמן המחזור של התדר היסודי של הדבור.

6.4 השפעת חלון Hann

הפעלת העבוד על בלוקים של N דגימות של אות הדבור הנה שוות ערך להכפלת סדרת הדגימות של האות בפונקציה זמנית $w_S(t)$ אשר נחן לתארה כך:

$$w_S(t) = \begin{cases} 1 & 0 \leq t \leq T_B \\ 0 & \text{אחרת} \end{cases} \quad (15)$$

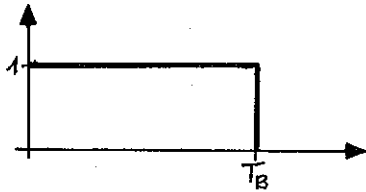
כאשר T_B הוא משך חלון הזמן ושוה לאורך הבלוק הקשור ל-N ע"י משוואה (11). הכפלה זו במשור הזמן מתקבלת כקונבולוציה, במשור התדר, בין ההתמרה של סדרת הדגימות לבין ההתמרה של

פונקציית החלון. ההתמרה של פונקציית החלון המלבני היא בעלת צורה של $\sin x/x$. צורה זו מצטיינת באונות צד (side lobes) גבוהות, כאשר הראשונה היא בערך חמישית מהאונה הראשית. אונות צד אלה עלולות לעוות את צורת הספקטרום ועל כן לפגוע (אולי) בבחירה הנכונה של הרכיבים הדומיננטיים לפי הקריטריונים של שיאים מקומיים. כמון כן במקרה של חלון מלבני יש חשש מאירציפיות במעבר מבלוק לבלוק בדבור המשוחרר. כדי להתגבר על כך אפשר לישים חלון זמן אחר. במשור החדר מטרת הישום של חלון כזה תהיה דיכוי אונות הצד ואילו במשור הזמן תהיה המטרה החלקת אירציפיות פוטנציאליות במעבר בין בלוקים באות המשוחרר.

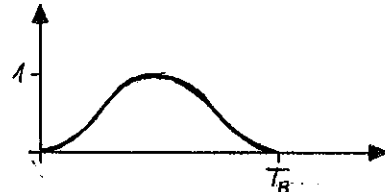
אחד החלונות מסוג זה הוא חלון Hann אשר הפונקציית הזמנית שלו $w_H(t)$ היא:

$$w_H(t) = \begin{cases} \frac{1}{2} \left(1 - \cos \frac{2\pi t}{T_B} \right) & 0 \leq t \leq T_B \\ 0 & \text{אחרת} \end{cases} \quad (16)$$

והיא מתוארת גם בצירור מס' 11.



א. חלון מלבני Rectangular Window



ב. חלון Hann Hann Window

ציור מס. 11: צורת "חלון זמן"

Fig. 11. Time Window shape

התמרת פוריה של פונקציית זו $w_H(\omega)$ תהיה:

$$W_H(\omega) = \frac{1}{2} 2\pi\delta(\omega) - \frac{1}{4} 2\pi\delta(\omega-\Omega) - \frac{1}{4} 2\pi\delta(\omega+\Omega) \quad (17)$$

$$\frac{2\pi}{T_B} = \Omega \quad \text{כאשר:}$$

ו- $\delta(\omega)$ הנה פונקציית דירק באפס.

כלומר ניתן לישים בקלות חלון זה במשור החדר (Otnes [18]). בתנאים הבאים:

$$N = 512, n = 16, r = 6, 1$$

הוכנס חלון Hann במשור החדר, לפני הצמצום, ע"י קונבולוציה של התמרת האות עם התמרת

חלון Hann . המובנות המתקבלת בנסוי זה (נסוי מס' 10) הושואה למובנות המתקבלת בנסוי שנערך בחנאים זהים אך עם חלון מלבני (נסוי מס' 5). התוצאות מופיעות בטבלה מס' 4.

טבלה 4: השפעת חלון זמן

Table 4. Time Window influence

סוג החלון Window type	מלבני	Hann
מובנות I(%) Intelligibility	73.9	72

כפי שרואים מן הטבלה הרי הכנסת חלון Hann גורמת דוקא לירידת המובנות, אם כי לא במידה ניכרת. הסיבה לכך יכולה להיות שהכפלה בחלון Hann מתקבלת כאפנון, של הדבור המשוחזר, באות סינוס המתאר את החלון ואפנון זה גורם לירידת המובנות. מחברר כמוכן מתוך הסתכלות בהדפסות של הצורה הזמנית של האות המשוחזר, שיש אירציפות קטנה במעבר מבלוק לבלוק במקרה של חלון מלבני, ועל כן לא הושג כל שפור בחחום זה ע"י חלון Hann .

כדאי לציין שהרכיבים שנבחרו בנסוי עם חלון זה היו שונים מן הרכיבים שנבחרו בנסוי עם חלון מלבני. היות והכנסת חלון Hann גם לא הביאה לשפור הטבעיות של הדבור המשוחזר אין כל סיבה להשתמש בחלון זה ועדיף חלון מלבני.

6.5 קצב ההעברה וחלות המובנות בו

לאחר הצמצום המתבצע במפריד יש להעביר אל המרכיב את המידע המתאר את האות המצומצם. בשיטת השיאים המקומיים יש להעביר לגבי כל אחד מ-n הרכיבים הנבחרים את מקומו במשור החדר (מספרו הסדורי), את המשרעת של החלק הממשי והמשרעת של החלק המדומה. נתונים אלה יש להעביר עבור כל בלוק מעובד.

מספר הרכיבים בכל בלוק הוא N , כאשר:

$$N = 2^a \tag{18}$$

כלומר נחוצות לכאורה a סביות לתאור מקומו של הרכיב, אך יש לזכור שאין צורך להעביר כל מידע על רכיבי הכבואה ויש על כן לחאר רק N/2 רכיבים, כלומר נחוצות רק a-1 סביות. בנוסף לכך במקום לתאר את מקומו של כל רכיב באופן מוחלט, אפשר לחאר את מקומו ביחס למקום

הרכיב הקודם, כלומר היאור אינקרמנטלי של המקום. נוסף לכך רק הרכיבים המתארים חחום התדרים 0.3 ק"ה עד 3 ק"ה הם המועמדים לבחירה היות ורוחב הסרט של הדבור צומצם לתחום זה, ועל כן עלינו לתאר רק 67.5 אחוז מחוך $N/2$ המקומות. היאור אינקרמנטלי של המקום והטעון האחרון מצביעים על כך שאפשר לחסוך סביח נוספת וס"ה להקציב 2-a סביות לתאור מקום הרכיב.

לתאור המשרעת של כל רכיב נחוצות $(r + 1)$ 2 סביות היות ולכל רכיב יש חלק ממשי וחלק מדומה המתוארים כל אחד ע"י r סביות למשרעת ועוד סביח אחת סימן. על כן כדי לתאר בלוק אחד של הדבור המצומצם בשיטת השיאים המקומיים יש צורך ב:

$$TR = n [(a - 2) + 2(r + 1)] \cdot \frac{fs}{N} \quad (19)$$

קצב ההעברה TR בין המפריד למרכיב ינתן אם כך ע"י הנוסחה:

$$TR = n(a + 2r) \frac{fs}{N} \quad (19א')$$

ולאחר פתיחת סוגריים פנימיים נקבל: כאשר fs הוא תדר הדגימה (8 ק"ה בנסויים שלנו). קצב ההעברה חושב בכל הנסויים והחוצאות רוכזו בטבלה 5.2 בנספת ג'. נביא כאן לדוגמה את קצבי ההעברה במספר נסויים, המיצגים את הקצב הנדרש בשיטת שיאים מקומיים.

טבלה 5: דוגמאות של קצבי העברה

Table 5. Samples of transmission rates

N	n	r	TR (bits/sec)
512	16	6	5250
256	8	6	5000
128	4	6	4750
128	4	4	3750
128	5	4	4687

קצב ההעברה יחסי ישר למספר הרכיבים הנבחרים, למספר סביות קוונטיזציה ולמספר הסביות לתאור המקום. החלות באורך הבלוק מסובכת יותר. מחוך (18) ניתן לחלץ את a כפונקציה של N ולהציב זאח ל-(19א') ואז נקבל:

$$TR = \frac{n}{N} fs \left(\frac{\log N}{\log 2} + 2r \right) \quad (19ב')$$

מכאן רואים שהגדלת N תגרום אמנם לירידת קצב ההעברה אך זה בהנחה ש- n איננו תלוי ב- N .
 אך נכון יותר יהיה להניח ש- n ו- N קשורים, בדרך כלל n/N גודל קבוע) ואז יגדל קצב ההעברה
 לוגריתמית עם עליית N . מכל האמור לעיל נובע שכדי לצמצם יותר את קצב ההעברה יש להקטין אחד
 או יותר מן הגורמים הבאים: מספר הרכיבים הנבחרים, מספר סביות קוונטיזציה ומספר סביות
 לתאור מקום הרכיב. נבדוק למשל בחנאים של

$$N = 512 \quad n = 16 \quad r = 4$$

באיזו מידה ניתן להקטין את קצב ההעברה ע"י שנוי r או n ובאיזו מידה תפגם המובנות.
 הקטנה r בסביבת אחת הקטין את הקצב בכ-11 אחוז והמובנות חרד בכ-8 אחוז. הקטנת n ב-2
 הקטין גם היא את הקצב בכ-11 אחוז והמובנות חרד בכ-4 אחוז.

הקשר בין המובנות לקצב ההעברה איננו פשוט אם כי באופן כללי ניתן לאמר שדרישה
 לקצב העברה נמוך מנוגדת לדרישה למובנות גבוהה. בציור מס' C.1 בנספח ג' תואר הקשר בין
 המובנות לקצב ההעברה עבור מרבית הנסויים שנערכו במשך המחקר. מתוך גרף זה ניתן לקבל
 מושג על הצירופים האפשריים של קצב העברה ומובנות בשיטת השיאים המקומיים בצרופיה השונים.
 בקצבי ההעברה של 4500 עד 5500 סביות לשניה ניתן להשיג מובנות של 70 אחוז ומעלה
 (להברוח). בקצב העברה של כ-2600 סביות לשניה ניתן להשיג מובנות של כ-55 אחוז. קצבי
 העברה מקובלים הם 2400, 4800 ו-9600 סביות לשניה. ניתן להשיג קצבים אלה בשיטת השיאים
 המקומיים עם מובנויות של 50, 70 ו-80 אחוז בהתאמה.

6.6 השפעת שנויים בשיטת הצמצום

6.6.1 תוספת קו שלישי של שפוע חיובי מתמשך

בשיטת השיאים המקומיים מיוצגת כל קבוצת קוים סביב שיא מקומי של פונקציית ספקטרום
 המשרעת ע"י הרכיב בנקודת השיא. אין בחירה זו מתחשבת במספר הקוים הנמצאים סביב השיא או
 במלים אחרות אין הבחירה מתחשבת על פני כמה קוים עולה הפונקציה מונוטונית עד השיא. השיטה
 הבאה מתחשבת בגורם זה בדרך הבאה: כל קו שלישי על קטע מונוטוני עולה ואשר איננו סמוך
 לקו שיא מקומי יחשב גם הוא כמועמד לבחירה, וזאת בנוסף לשיאים מקומיים. הקוים הגדולים
 ביותר מבין המועמדים משני הסוגים יבחרו כמיצגים את הבלוק של ההחמרה. להדגמת שיטת בחירה
 זו נחבונן שוב בציור מס' 5ב'. בציור זה מסומנים הקוים המהוים קו שלישי על שפוע חיובי
 ע"י חץ. כפי שרואים, הרי במקרה זה רוב הקוים האלה מהווה ממילא שיא מקומי ועל כן היה
 מועמד גם קודם. אך קו מס' 24 הוא קו שלישי על שפוע עולה והוא איננו שיא מקומי ואף לא
 סמוך לו. שמונה הקוים שיבחרו בשיטה זו (שיאים מקומיים בתוספת כל קו שלישי על שפוע חיובי)
 סומנו ברבוע (קוים מס' 6, 8, 12, 16, 24, 26, 28, 43). כפי שרואים ישנו שנוי קטן בבחירת
 הקוים, אזור הקוים 22 עד 27 מיוצג כעת ע"י שני קוים (24 ו-26) במקום ע"י קו אחד קודם (26).
 בחנאים הבאים:

$$N = 256, \quad n = 9, \quad r = 6, 4$$

נערך נסוי אשר בו בבחירת הרכיבים הדומיננטיים החשבו ברכיבים המיצגים קו שלישי על שפוע חיובי (נסוי מס' 22). נבדקה השפעת בחירה זו על המובנות בהשוואה לשיטת הבחירה הרגילה של שיאים מקומיים (נסוי מס' 24). התוצאות מופיעות בטבלה מס' 6 להלן:

טבלה 6: השפעת חוספת קו שלישי על שפוע חיובי

Table 6. Influence of third line on positive slope

שיטת בחירה	שיאים מקומיים בחוספת שלישי על שפוע חיובי	שיאים מקומיים בלבד
מובנות I (%)	69.3	73.2

מתברר מן התוצאות שהתחשבות כזו ברוחבו של השיא המקומי פוגעת במובנות היותן והמובנות בשיטה זו נמוכה בכ-4 אחוזים מן המובנות בשיטת השיאים המקומיים. מעניין לציין שמתוך ההדפסה של הרכיבים הנבחרים ומתוך חשוב אחוז הקוים הנבחרים באזורי חדר שונים מתברר שבשיטת קו שלישי נבחרים יותר רכיבים בחדרים הנמוכים ופחות רכיבים בחדרים הגבוהים כלומר מודגש למעשה הפורמנט הראשון וזו אולי הסיבה לירידת המובנות. ברור אם כך שדוקא הרכיבים המיצגים את השיאים המקומיים של פונקצית הספקטרום הם הדומיננטיים לצורך צמצום.

6.6.2 אפיון רכיב סמוך לנבחר ע"י סביבת אחת

במטרה לשפר את מובנות הדבור המשוחזר המחקבל עם צמצום מספר הרכיבים ל-1/32 וזאת

מבלי להגדיל במדה ניכרת את קצב ההעברה נבדק הרעיון הבא:

במקום לאפס את כל הרכיבים אשר אינם נבחרים יאופסו רק כל אותם הרכיבים אשר אינם סמוכים לרכיבים הנבחרים. רכיב סמוך לרכיב נבחר הוא הרכיב המופיע לפני או אחרי הנבחר, כלומר מספרו עוקב למספר הרכיב הנבחר. כל אחד מן הרכיבים הסמוכים לנבחר יחזאר ע"י סביבת אחת בלבד (החלק הממשי והמדומה גם יחד), אשר תהיה למעשה קוונטיזציה של סביבת אחת של משרעת של כל חלק של הרכיב ביחס למשרעת של החלק המתאים ברכיב הנבחר. הסימן של החלק הממשי של הרכיב הסמוך יהיה זהה לסימן החלק הממשי של הרכיב הנבחר וכן בהתאמה לחלק המדומה. הקריטריון לקביעת הסביבת יהיה מבוסס על רמת האנרגיה הנישאת ע"י הרכיב הסמוך ביחס לרמת האנרגיה ברכיב הנבחר. נסמן ב-1 את מספר הרכיב הנבחר, ב- $A(1)$ את הרכיב של ההתמרה וב- $S_A(1)$ את המשרעת של הקו הספקטרילי המתאים, אזי הסביבת הנוספת $b_{0,1}$ תקבע כך:

$$b_{0,1} = \begin{cases} 1 & S_A(1 \pm 1) > S_A(1)/4 \\ 0 & \text{אחרת} \end{cases} \quad (20)$$

כאשר b_0 הסבית עבור הרכיב לפני הנבחר ו- b_1 הסבית עבור הרכיב אחרי הנבחר. העברת סבית זו מגדילה בשתיים את מספר הסביות שיש להעביר עבור כל רכיב נבחר. במרכיב יקבע הערך של

החלק הממשי והמדומה של הרכיב הסמוך בדרך הבאה:

$$\begin{aligned} \operatorname{Re} \{A(1 \pm 1)\} &= \begin{cases} \operatorname{Re} \{A(1)\} / 2 & b_{0,1} = 1 \\ 0 & b_{0,1} = 0 \end{cases} \\ \operatorname{Im} \{A(1 \pm 1)\} &= \begin{cases} \operatorname{Im} \{A(1)\} / 2 & b_{0,1} = 1 \\ 0 & b_{0,1} = 0 \end{cases} \end{aligned} \quad (21)$$

בחנאים הבאים:

$$N = 512, n = 16, r = 6, 6$$

נערך נסוי אשר בו חוארו הרכיבים הסמוכים לנבחרים בשיטה אשר חוארה לעיל (נסוי מס' 34). נבדקה ההשפעה על המובנות של חוספת שחי סביות אלה בהשוואה למובנות המושגת בנסוי שנערך באותם חנאים אך בו כל הרכיבים הלא נבחרים אופסו (נסוי מס' 33). החוצאות מופיעות בטבלה מס' 7.

טבלה 7: השפעה חוספת סבית לחאור רכיב סמוך

Table 7. Influence of additional bit for adjacent component

חאור הרכיבים הסמוכים לנבחרים	מאופינים ע"י סבית אחת כ"א	מאופסים
I(%) מובנות	54.3	58.9
קצב העברה TR(bits/sec)	5750	5250

מן החוצאות מחברר שחוספת שחי הסביות בשיטה המחוארה לעיל גורמת דוקא לירידה המובנות בכ-4 אחוז. הסיבה לכך יכולה להיות הקוונטיזציה הגסה מדי וכן השגיאה המתקבלת במופע כחוצאה מהשוואת הסימן של החלק הממשי והמדומה של הרכיב הסמוך לאלה של הרכיב הנבחר מה שכרוך לעתים בהחלפת הסימן ועל כן בשגיאה גסה במופע. היות ובשיטה זו קצב ההעברה גדל ואילו המובנות יורדת אין טעם להשתמש בה.

6.6.3 קוונטיזציה במספר סביות מודרג

Campanella [3] מציין במאמרו שהשונות של רכיבי התמרת פוריה שונה מרכיב לרכיב ובאופן כללי היא יורדת עם עליה סדר הרכיב (שוה ערך לעליה התדר שהרכיב הספקטרולי מייצג). על כן ניתן לבצע קוונטיזציה של כל רכיב במספר רמות המתאים לשונות שלו, תוך שמירה על אותה שגיאה, כלומר לרכיבים אשר השונות שלהם קטנה אפשר להקצות פחות סביות בקוונטיזציה. אמנם הנאמר ב-[3] מתייחס לקוונטיזציה לנאליה אך ניסינו לישם זאת גם בקוונטיזציה לוגריתמית. ניתן לחלק את תחום התדרים לשלושה תחומים אשר בכל אחד מהם השונות של הרכיבים פחות או יותר זהה ועל כן מספר הסביות שיוקצו לכל הרכיבים בתחום כזה יהיה זהה. שלושת התחומים הם: א. 0.3 ק"ה עד 1 ק"ה, ב. 1 ק"ה עד 2 ק"ה, ג. 2 ק"ה עד 3 ק"ה. אם בתחום א' נקצה לכל רכיב i סביות לקוונטיזציה הרי בתחום ב' נוכל להקצות $i-1$ סביות ואילו בתחום ג' נוכל להקצות $i-2$ סביות.

הרעיון הוא אם כך לנסות ולהקטין את קצב ההעברה בתנאים בהם נותרו $1/32$ מן הרכיבים ע"י צמצום נוסף שיושג בעזרת הקצבה מודרגת של סביות קוונטיזציה ולבחון באיזה מדה תפגע המובנות.

כדי להעריך את הצמצום שניתן להשיג בדרך זו כלומר בכמה יקטן קצב ההעברה יש לדעת איזה חלק מ- n הרכיבים הנבחרים נבחר מכל אחד משלושת התחומים. לצורך זה, תוך כדי העבוד בנסויים השונים, חושב והודפס אחוז הרכיבים הנבחרים לפי שלושת התחומים. לאחר מכן חושב הממוצע באחוזים, של הקוים הנבחרים לפי שלושת תחומי החדר. תוצאות אלה רוכזו בטבלה C.1 בנספח ג'. בטבלה מס' 8 מופיע הממוצע על פני כל הנסויים.

טבלה 8: אחוז הקוים הנבחרים בממוצע בכל תחום

Table 8. Average percent of lines chosen in each band

i	0	1	2
תחום תדרים (ק"ה)	1-0.3	2-1	3-2
אחוז קוים נבחרים $P(i)$	68	24.7	7.3

מספר הקווים הנבחרים בכל אזור $n(i)$ נחון ע"י

$$n(i) = n \cdot P(i)$$

כאשר $P(i)$ הוא אחוז הקווים הנבחרים בכל אזור. קצב ההעברה יהיה:

$$TR = \left[n(a-2) + 2 \sum_{i=0}^2 n(i)(r-i+1) \right] \frac{fs}{N} =$$

$$= n \left[(a-2) + 2 \sum_{i=0}^2 P(i)(r-i+1) \right] \frac{fs}{N} \quad (19ג')$$

ולפיכך עבור $a = 9, r = 6$

יהיה החסכון בקצב ההעברה כ-4 אחוז. בתנאים הבאים:

$$N = 512, n = 16, r = 6, 6$$

נערך נסוי בו נוסתה קוונטיזציה במספר סביות מודרג כאשר בתחום א' הוקצו 6 סביות, בתחום ב' הוקצו 5 סביות ובתחום ג' הוקצו 4 סביות (נסוי מס' 35). נבדקה המובנות בנסוי זה בהשוואה למובנות המתקבלת באותם התנאים ובאשר הקוונטיזציה קבועה בכל התחומים לשש סביות (נסוי מס' 33). גשני המקרים הבחירה היתה בשיטת שיאים מקומיים. החוצאות מופיעות בטבלה מס' 9.

טבלה 9: השפעה קוונטיזציה מודרג

Table 9. Influence of staggered quantization

שיטת קוונטיזציה	מודרג	רגילה
מובנות I(%)	54.7	58.9
קצב העברה TR(bits/sec)	5050	5250

ניתן בשיטה זו להקטין את קצבהעברה בכ-4 אחוז אך בתנאי שמוכנים ליותר על 4 אחוז של המובנות.

6.7 איכות הדבור המשוחרר

משלושת הקריטריונים המגדירים את איכות הדבור המשוחרר תואר עד עתה רק אחד והוא המובנות. נסכם רק שבשיטת שיאים מקומיים ניתן להשיג מובנות של 70 עד 90 אחוז עבור הברות במלים. נסווג להלן את הטבעיות ונתאר במדח מה כיצד נשמע הדבור המשוחרר.

בשיטת שיאים מקומיים כאשר מותירים $1/32$ מרכיבי ההחמרה בלבד ובארכי בלוקים של 16 עד 64 מילישניות מתקבל דבור אשר ניתן לסווג את טבעיותו כמספקת עד טובה בסולם של 5 דרגות כפי שהוצג בסעיף 3.4.2 [1]. הדבור המשוחרר כולל לעתים צלילי לואי אקראיים הנשמעים כצלילי כסילופון. הדבור המשוחרר כולל גם "Hum" עם הרמוניות שלו, הנובע מן הממיר הספרתי לאנלוגי של המחשב ואשר אותו לא ניתן היה לסלק. נתאר בקצרה את טבעיות הדבור המשוחרר בתלות בפרמטרים השונים. בבלוקים של 128 מילישניות נשמע הדבור כמתמשך בצורה לא טבעית. עם הקטנת אורך הבלוק ל-64 ו-32 מילישניות (תוך שמירה על $n/N = 1/32$) משתפרת הטבעיות. בבלוקים קצרים מאד (8 מילישניות ופחות) מתגברים מאד צלילי הלואי וגורמים להרגשת דבור לא טבעי.

הקטנת מספר הרכיבים הנבחרים מוסיפה צלילים זרים בכמות הולכת וגדלה. העלאה מספר הרכיבים הנבחרים ל- $1/16$ משפרת את טבעיות הדבור כך שניתן להגדירה כטובה.

להקטנת מספר סביות קוונטיזציה השפעה קטנה על הטבעיות בתחום רחב של שנוי. קשה מאד להבחין בהבדל בין 8 ל-4 סביות. עם זאת צמצום נוסף במספר הסביות גורם להופעת עוותים וצלילי לואי רבים. הקריטריון השלישי להערכת איכות דבור הוא אפשרות זהוי הדובר. מעט ניתן לאמר בנקודה זו ביחס לשיטת שיאים מקומיים היות ובכל הנסויים היה אותו דובר. לא נערכה בדיקה מקיפה בנקודה זו היות ובדיקה כזו לא הוגדרה במטרות העבודה. עם זאת יש לציין שמרבית המאזינים זיהו את הדובר בנסויים בהם היו הטבעיות והמובנות טובים.

6.8 חוסר תלות בין מובנות הדבור המשוחרר לקרוסקורלציה ושגיאה רבועית ממוצעת

הערכת שיטת עבוד של אות הדבור המיועדת לצמצם את תכולתו, מתבססת בדרך כלל על המובנות של הדבור המשוחרר לאחר צמצום. מדידת המובנות כרוכה בבדיקות ממושכות ומיגעות עם מספר רב של מאזינים (כפי שזה תואר בסעיף 6.2). במדה והעבוד המוצע מדומה במחשב ספרתי (טכניקה שהיא מקובלת מאד היום) הרי הצורך בבדיקות מובנות מחייב שחזור מלא של אות הדבור עד כדי הפיכתו לאות אנלוגי והקלטתו לצורך הבדיקות. דרך זו של בצוע מחקר, מיגעת ודורשת זמן רב. לו ניתן היה למצא התאמה בין בטוי אנליטי כלשהו לבין המובנות, כפי שהיא נמדדת בבדיקות עם מאזינים, הרי ניתן היה לפשט את ההליכי המחקר ע"י שלוב שתי הפעולות של העבוד וההערכה במחשב ספרתי. יובא פה תאור של החשובים שנוסו למרות שנכשל

הנסיון למצא התאמה בינם לבין המובנות כפי שהתקבלה מנסויי המובנות. הסיבה לכשלוך נעוצה כנראה בעובדה שהמובנות היא תופעה מורכבת מאד אשר קשה להעריכה ע"י חשבים אנליטיים פשוטים יהסית.

כזכור, העבוד בוצע על בלוקים שכללו N דגימות כ"א. בכל בלוק חושבה הקרוסקורלציה באפס, במשור הזמן, בין אות הדבור המקורי לבין אות הדבור המשוחזר (לאחר העבוד). הקרוסקורלציה של כל הקטע שעבר עבוד חושבה כממוצע הערכים שהתקבלו עבור כ"א מן הבלוקים. נציין כי תחילה חושבה פונקציה הקרוסקורלציה על פני בלוק שלם ($-N$ עד $+N$) ונמצא שהמכסימום של הפונקציה מתקבל באפס, כלומר אין העבוד מכניס השחיה.

נסמן ב- $X(k)$ את הדגימה ה- k בחוך בלוק m של אות הדבור המקורי. וב- $Y(k)$

את הדגימה ה- k של אות הדבור המשוחזר. אזי הקרוסקורלציה המנורמלת באפס תהיה נתונה ע"י

$$R_{xy}(0)_m = \frac{\sum_{k=0}^N X(k) Y(k)}{\left(\sum_{k=0}^N X(k)^2 \sum_{k=0}^N Y(k)^2 \right)^{\frac{1}{2}}} \quad (22)$$

הנרמול נעשה כדי למנוע חלות בגודל האות. הקרוסקורלציה של כל הקטע חושבה כממוצע הערכים בכל בלוק:

$$R_{xy}(0) = \frac{1}{M} \sum_{m=1}^M R_{xy}(0)_m \quad (23)$$

כאשר M הוא מספר הבלוקים בקטע הדבור שעבר עבוד. במקביל לחשוב הקרוסקורלציה חושבה גם השגיאה הרבועית הממוצעת בין האות המקורי לבין האות המעובד. השגיאה MSE_m בחוך בלוק m נתונה ע"י:

$$MSE_m = \frac{\sum_{k=0}^N (X(k) - Y(k))^2}{\sum_{k=0}^N X(k)^2} \quad (24)$$

גם כאן נעשה נרמול למניעת חלות בגודל האות. השגיאה בכל קטע חושבה כממוצע הערכים בכל בלוק:

$$MSE = \frac{1}{M} \sum_{m=1}^M MSE_m \quad (25)$$

תוצאות חשובים אלה כפי שהודפסו בתום כל עבוד רוכזו בטבלה 3. C בנספח ג'. לאותו נספח צורפו גם ציורים C.2 ו-C.3 המתארים בצורה גרפית את הקשר בין המובנות לבין הקרוסקורלציה והשגיאה הרבועית הממוצעת.

מתברר שהחוצאות חלויות בטקסט ואפילו עבור אותו טקסט אין התאמה בין המובנות לבין אחד משני הבטויים האלה. אחת הסיבות האפשריות לכשלון ההתאמה קשורה לעובדה שקטע דבור מתחלק לקטעים הכוללים מלים ולקטעים שקטים שהם הפסקות בין מלים (או הברות). התברר מתוך הדפטי האות הזמני, לפני ואחרי העבוד, שקימת התאמה טובה בין שניהם כאשר האות חזק (קימת מלה) והתאמה גרועה כאשר האות חלש. ההתאמה הגרועה בהפסקות בין המלים אינה פוגעת במובנות. הקרוסקורלציה או השגיאה חושבו כממוצעים על פני כל הבלוקים בכך התחשבו גם בשגיאות בהפסקות. עדיף היה לבצע את המיצוע רק על קטעים שכללו מילים ממש. אך חשוב כזה הרבה יותר מסובך ולא בוצע במחקר זה.

7. סיכום והמלצות

מטרת המחקר היתה לפתח שיטה לצמצום תכולת אות הדבור אשר לא תסחמך על מדידת הפרמטרים של פונקציית הערור, אך עם זאת תבטיח דבור מובן וטבעי עם צמצום ביחס של 1:10 לפחות. שיטת השיאים המקומיים שהוצעה מקימה דרישות אלה עם בחירה נכונה של הפרמטרים הקובעים את טיבה. נקבעה הנקודה בה מחקבל ניצול אופטימלי של השיטה וכך הובררה התלות של איכות הדבור המשוחזר בפרמטרים השונים.

7.1 סיכום האפשרויות האופטימליות לניצול שיטת השיאים המקומיים

7.1.1 נקודת עבודה אופטימלית

השיטה תוארה בסעיף 5.3.1, כאשר נדגיש שוב שהבחירה של רכיבי החמרת פוריה הדיסקרטי מבוססת על שיאים מקומיים של פונקציית הספקטרום הדיסקרטי והקוונטיזציה של הרכיבים הנבחרים הנה לוגריתמית. נקודת העבודה האופטימלית במרחב המוגדר ע"י שלושת הפרמטרים הבסיסיים הנה:

- א. בצוע החמרה בבלוקים שארכם 16 מילישניות והכוללים 128 דגימות.
- ב. בחירת $1/32$ מרכיבי החמרה ואפוס כל האחרים, דבר המתבטא בבחירת 4 רכיבים בלבד בבלוק האופטימלי.
- ג. קוונטיזציה של משרעת הרכיבים הנבחרים עם 4 סביות.

בחנאים אלה קצב ההעברה הנדרש הוא 3750 סביות לשניה, אשר מבטא צמצום אות הדבור ביחס של 1:15. הדבור המתקבל מצטיין במובנות גבוהה של 70 עד 90 אחוז של הברות בתוך מלים. טבעיות הצליל המתקבל מדורגת כמספקת עד טובה, כאשר הדבור מלווה לעתים בטונים אקראיים המתנגנים ברקע.

7.1.2 אפשרויות ניצול קצבי העברה סטנדרטיים

כידוע קימים קצבי העברה, סטנדרטיים להעברת נתונים ספרתיים. שלושה מן הקצבים האלה הם: 9600, 4800 ו-2400 סביות לשניה. בשיטת שיאים מקומיים ניתן להשתמש בקצבים אלה בהרכבים שונים של הפרמטרים. נציין להבא הרכבים אלה בדרך הבאה: $N/n/x$

בקצב העברה של 9600 סביות לשניה ניתן לקבל דבור בעל מובנות טובה יותר מאשר בנקודת האופטימום וזאת ע"י העלאת מספר הרכיבים הנבחרים ומספר הסביות בקוונטיזציה. נקבל מובנות של 80 אחוז לפחות וטבעיות טובה ללא כל צלילי לוואי, להלן מספר אפשרויות מיצגות: $128/10/4$, $128/9/5$, $256/16/5$, $512/31/5$.

בקצב העברה של 4800 סביות לשניה ניתן לקבל דבור בעל מובנות מעט טובה יותר מאשר
בנקודת האופטימום, כלומר 75 אחוז לפחות וכן תשופר במידת מה הטבעיות. האפשרויות הן:
128/5/4, 128/4/6, 256/9/4, 256/8/5.

בקצב העברה של 2400 סביות לשניה ניתן לקבל מובנות של כ-55 אחוז וטבעיות ירודה
למדי עם החווספות ניכרת של טונים אקראיים. בחנאים אלה נוסחה אמנם רק האפשרות 512/8/5
אך נראה שהרכבים כמו 128/2/6, 256/4/5 יתנו חוצאות דומות אם לא טובות יותר.

7.2 המלצות להמשך

נציין אפשרויות שונות שכדאי לחקור בעתיד כהמשך לעבודה זו ואשר עשויות לשפר את
איכות הדבור המשוחרר או לצמצם צמצום נוסף את קצב ההעברה הנחוץ.

- א. נצול הקורלציה הקימת בין רכיבי ההתמרה בבלוקים סמוכים. ניצול קורלציה זו יכל להתבטא
במסירת מיקום ואולי אף משרעת של הרכיבים הנבחרים יחסית לרכיבים שנבחרו בבלוק הקודם.
- ב. נצול הקורלציה הקימת בין רכיבי התמרת פוריה של אות הדבור בתוך בלוק אחד, ע"י
בצוע התמרה לנאריה של הרכיבים ובחירה של הרכיבים הדומיננטיים לאחר התמרה זו.
רכיבי התמרת פוריה, ישוחררו במרכיב ע"י התמרה לנאריה שניה ולאחריה ימשך תהליך
השחזור כפי שתואר בשיטת השיאים המקומיים. שיטה זו דומה להצעתם של Crowther and
Rader [6] עבור ווקודר ערוצי.
- ג. מדידת המשרעת המקסימלית של הרכיבים בכל בלוק ובצוע הקוונטיזציה ביחס לרמה זו.
התאמת הרמה בכל בלוק תאפשר צמצום נוסף במספר הסביות לקוונטיזציה ותדרוש רק תוספת
של מספר קטן של סביות בכל בלוק, לתאור הרמה המקסימלית.
- ד. העברת הנתונים על הרכיבים הנבחרים בקואורדינטיות קוטביות כלומר ע"י משרעת ומופע.
תבוצע קוונטיזציה לוגריתמית של המשרעת וקוונטיזציה לנאריה של המופע. הקוונטיזציה
של המופע תהיה עם שתי סביות או לכל היותר שלוש סביות.

8. נ ס פ ח י ם

נספח א' - פרטי הנסויים ותוצאות נסויי מובנות

בנספח זה בטבלה A.1 מופיעים נתונים המחארים את הנסויים השונים.
הנסויים מוספרו וההתייחסות אליהם בטקסט היא לפי מספרים אלה.

בטבלה A.2 מרוכזות תוצאות נסויי מובנות בנסויים השונים
כשהן מפורטות לפי מאזינים, וכן הממוצע בכל נסוי.

Table A.1 Details of tests

Test No.	3	4	5	6	7	8	10	12	13	14	15	16	17	19	20
מס' דגימות לבלוק No. of samples per block	512	512	512	512	512	512	512	512	512	512	256	128	1024	256	128
מס' קווים נבחרים No. of chosen lines n	32	256	16	64	8	4	16	16	16	16	9	5	31	8	4
מס' מביינות בקוונטיזציה No. of quantization bits r	6	6	6	6	6	6	6	8	4	5	6	6	6	6	6
לוח מס' Text No.	1	1	1	1	1	1	1	3	3	3	3	3	3	4	4
שיטת בחירה Choosing method	שיטת מס' מקומיים														
צורת "חלון" זמן Time window shape	מלבני														
פרטים נוספים Additional details															

Table A.1 Details of tests (continuation)

פרטי הנסויים השונים (המשך) טבלה A.1

נסוי מס' Test No.	21	22	23	24	27	28	29	30	31	32	33	34	35	36	37
מספר דגימות לבקור No. of samples per block N	64	256	128	256	512	512	512	512	64	512	512	512	512	512	32
מספר קווים נבחרים No. of chosen lines n	2	9	5	9	16	16	16	16	2	12	16	16	16	16	1
מספר סביבות קווינטזציה No. of quantization bits r	6	6	6	6	6	6	6	6	2	6	6	6	4-6*	3	6
לוח מס' Text No.	4	4	4	4	5	5	3	3	3	1	6	6	6	3	3
שיטה בחירה Choosing method	שיאים מקומיים	שיאים מקומיים ושלישי על שפוע עולה	שיאים מקומיים												
צורת "חלון זמן" Time window shape	↔														
פרטים נוספים Additional details					כולל מידע של מביה אחת על הרכיב הסמור לנבחר								כולל מידע של מביה אחת על הרכיב הסמור לנבחר	* מספר סביבות קווינטזציה זרידה עם עליית התדר, בשני שלבים	

Table A.2 Results of Intelligibility tests

טבלה א.2 תוצאות נסיונות אינטליגיביליות

Test No.	3		4		5		6		7		8		10		12	
	Listener	Intellig. I (%)	Listener	Intellig. I (%)	Listener	Intellig. I (%)	Listener	Intellig. I (%)	Listener	Intellig. I (%)	Listener	Intellig. I (%)	Listener	Intellig. I (%)	Listener	Intellig. I (%)
1		82	3	95.5	1	77	14	85	10	41	10	13.5	16	77	10	81
5		82	4	95.5	2	77	19	77	11	59	11	36.5	17	64	20	70
6		82	5	95.5	3	82	21	87	13	32	12	9	18	68	13	73
8		82	7	95.5	5	77	22	82	19	50	13	18	19	64	3	85
9		64	8	100	8	64	1	77	63	59	69	9	20	77	2	85
34		77.5	9	95.5	23	64	56	77.5	64	59	70	5	75	82	10	88.5
35		73			33	68	57	87	65	59	71	18	76	54.5	81	73
36		73			32	73	58	91	66	73	72	5	77	73	82	73
38		77.5			30	73	59	91	67	59	73	18	78	77.5	83	69
42		82			28	64	60	87	68	54.5	74	14	79	82	84	77
47		91			26	77.5	61	87	99	54.5			80	73		
45		87			25	77.5	62	96	98	50						
46		82			24	87										
49		77.5														
48		73														
Total	15	1185.5	6	577.5	13	961	12	1024.5	12	650	10	146	11	792	10	774.5
Average		79		96.2		73.9		85.3		54.2		14.6		72		77.5

Table A.2 Results of Intelligibility tests (continuation)

(תוצאות מבחני ההבנת שמע A.2 המוצגים)

Test No.	13		14		15		16		17		19		20		21	
	Listeners	Intellig. I (%)	Listeners	Intellig. I (%)	Listeners	Intellig. I (%)	Listeners	Intellig. I (%)	Listeners	Intellig. I (%)	Listeners	Intellig. I (%)	Listeners	Intellig. I (%)	Listeners	Intellig. I (%)
	18	60	23	70	8	93.5	32	77	33	37	6	75	11	75	8	34
	16	70	19	77	31	90	30	80	8	30	13	62.5	30	69	2	37.5
	7	63	11	53	17	90	29	93.5	30	23	27	72	25	78	29	34
	6	73	1	53	26	97	28	100	32	27	21	59	16	75	36	46.5
	21	67	22	83	25	90	27	90	34	40	22	62.5	29	78	54	43.5
	41	70	24	70	24	87	21	87	40	40	76	63	37	72	59	37
	40	77	38	63	34	93.5	33	93.5	37	37	75	67	71	77	60	40
	85	63	20	67	35	93.5	36	90	20	20	77	67	72	67	61	60
	86	73	42	63	39	90	18	90	17	17	78	67	73	63	62	46.5
	87	63	89	67	31	93.5	92	93.5	23	23	80	67	74	63	70	46.5
			90	73	44	100					79	50			69	26.5
					43	90										
					19	90										
No. Total	10	669	11	739	13	1198	10	894.5	10	294	11	712	10	717	11	452
Average		66.9		67.1		92.1		89.5		29.4		64.7		71.7		44.1

Table A.2 Results of Intelligibility tests (combination)

(תוצאות מבחני הבנת שמע - תוצאות משולבות)

Test No.	22	23	24	27	28	29	30	31
77777 Listener	75	84	72	63	60	83	20	93.5
77777 Intellig- I (%)	66	84	72	43.5	46.5	70	30	87
77777 Listener	26	26	41	48	47	45	27	23
77777 Intellig- I (%)	59	62.5	81	40	63	70	36.6	21
77777 Listener	21	39	42	49	46	50	6	2
77777 Intellig- I (%)	69	75	65.5	46.5	50	67	13.3	93.5
77777 Listener	18	23	10	43.5	105	60	20	11
77777 Intellig- I (%)	75	75	65.5	43.5	46.5	60	13.3	90
77777 Listener	23	63	44	23	101	77	25	54
77777 Intellig- I (%)	70	70	73	33	43.5	60	43.5	93.5
77777 Listener	21	70	43	40	60	104	98	36
77777 Intellig- I (%)	67	77	73	40	40	63	26.5	77
77777 Listener	2	2	20	26.5	104	102	99	67
77777 Intellig- I (%)	63	83	70	30	46.5	57	26.5	83
77777 Listener	56	70	38	30	40	103	64	68
77777 Intellig- I (%)	70	70	70	17	40	63	10	87
77777 Listener	57	77	17	30	40	63	63	65
77777 Intellig- I (%)	67	80	80	17	40	63	26.5	93.5
77777 Listener	58	80	27	17	40	63	10	90
77777 Intellig- I (%)	80	80	80	17	40	63	26.5	87
77777 Listener	6	6	6	17	40	63	10	90
77777 Intellig- I (%)	80	80	80	17	40	63	26.5	87
77777 Listener	12	12	12	36.3	49.6	9	27	11
77777 Intellig- I (%)	900.5	900.5	952	36.3	49.6	607	269.7	981.5
77777 Listener	13	13	13	36.3	49.6	9	27	11
77777 Intellig- I (%)	73.2	73.2	73.2	36.3	49.6	67.4	27	89.2
77777 Listener	14	14	14	36.3	49.6	9	27	11
77777 Intellig- I (%)	971	971	971	36.3	49.6	607	269.7	981.5
77777 Listener	12	12	12	36.3	49.6	9	27	11
77777 Intellig- I (%)	69.3	69.3	69.3	36.3	49.6	67.4	27	89.2
77777 Listener	12	12	12	36.3	49.6	9	27	11
77777 Intellig- I (%)	73.2	73.2	73.2	36.3	49.6	67.4	27	89.2
77777 Listener	12	12	12	36.3	49.6	9	27	11
77777 Intellig- I (%)	69.3	69.3	69.3	36.3	49.6	67.4	27	89.2
77777 Listener	12	12	12	36.3	49.6	9	27	11
77777 Intellig- I (%)	69.3	69.3	69.3	36.3	49.6	67.4	27	89.2

Table A.2 Results of Intelligibility tests (continuation)

טבלה A.2 תוצאות מבחני הבנת שמע (המשך)

Test No.	32		33		34		35		36		37	
	Listener	Intellig. I (%)	Listener	Intellig. I (%)	Listener	Intellig. I (%)	Listener	Intellig. I (%)	Listener	Intellig. I (%)	Listener	Intellig. I (%)
	24	68	6	70	24	53	21	60	76	67	55	57
	26	64	20	67	26	57	23	73	75	67	58	57
	52	54.5	17	73	52	63	2	60	79	46.5	57	46.5
	43	50	27	63	43	57	11	63	80	53	59	73
	23	64	38	90	76	53	69	43.5	78	67	60	70
	21	64	56	43.5	75	46.5	70	40	77	73	62	60
	11	77.5	58	30	80	50	72	70	73	53	61	53
	2	87	57	53	79	50	71	57	74	43.5	70	63
	36	59	59	46.5	78	57	73	37	71	67	69	50
	54	64	60	43.5	77	57	74	43.5	72	50	56	73
			61	57								
			62	70								
# of Total	10	652	12	706.5	10	543.5	10	547	10	587	10	602.5
Average		65.2		58.9		54.3		54.7		58.7		60.2

נספח ב' - לוחות המלים אשר שמשו לנסויים

לוח מס'					
6	5	4	3	1	
שמך	מפה	אדום	דלק	שקע	1
יונה	גדר	הגר	מספיק	מורה	2
דבק	חזק	קדש	פתל	יובל	3
קטר	ימה	טיפוס	זאב	דגל	4
טיפוס	תפוז	דבק	חלום	תוכי	5
לבן	סופר	לבן	חמר	פילון	6
צבע	קטן	ערך	קומה	כלה	7
גבר	אגס	מופח	שמך	חתך	8
אדום	בירה	רימון	לוי	רופא	9
חלב	עגול	פקח	בורג	סוכן	10
קופה	כבד	מיטה	יונה	לחם	11
זמר	טליח	ספל	סולם		12
מיטה	משל	זעיר	ציבור		13
שלג	נערה	צלם	גנב		14
תודה	עצב	יחד	נבל		15

נספח ג' - ריכוז תוצאות שונות

1. אחוז הקוים הנבחרים לפי אזורים

בטבלה C.1 רוכזו תוצאות חשובות אחוז הקוים הנבחרים בכל אחד משלושת החומי התדרים: 0.3 עד 1 ק"ה, 1 עד 2 ק"ה, 2-3 ק"ה. התוצאות המופיעות הן עבור נסויים שונים אשר הנם בלתי תלויים לצורך זה, כלומר הם נבדלים זה מזה במספר הקוים הנבחרים או באורך הבלוק או בטקסט. התוצאות בכל נסוי הן ממוצע על פני כל הבלוקים של הקטע המעובד. בסוף הטבלה חושבה תוצאה ממוצעת.

2. קצב העברה בנסויים שונים

עבור כל הנסויים שבוצעו חושב קצב ההעברה. נתון זה נרשם בטבלה C.2 יחד עם מספר הנסוי והמובנות שנמדדה באותו נסוי. החשוב בוצע במרבית הנסויים לפי נוסחה (19א'). נסוי מס' 35 חושב קצב ההעברה לפי (19ג').

בנסויים מס' 27 ו-34 נוספות 2 סביות לכל רכיב לתאור הרכיבים הסמוכים ועל כן

הנוסחה היא:

$$TR = n (a + 2r + 2) \cdot \frac{fs}{N}$$

בנסוי 4 נבחרו כל הרכיבים ולכן אין כל צורך לתאר את מקומו של כל רכיב אלא מספיק לתאר את משרעתם של $N/2$ רכיבים. כמו כן רק הרכיבים בתחום התדרים 0.3 עד 3 ק"ה מיצגים את אות הדבור ועל כן במקרה זה חושב קצב ההעברה לפי:

$$TR = 2(r + 1) \frac{N}{2} \cdot \frac{2700}{4000} \cdot \frac{fs}{N} = (r + 1) \cdot fs \cdot \frac{27}{40}$$

תוצאות מרבית הנסויים תוארו גם בגרף בציור מס' C.1 המחאר את המובנות בתלוח בקצב ההעברה, כאשר הנקודות בגרף סומנו ע"י מספרי הנסוי בהם הם התקבלו.

3. קרוסקורלציה ושגיאה רבועית ממוצעת

בטבלה C.3 רוכזו תוצאות חשובי הקרוסקורלציה והשגיאה הרבועית הממוצעת כפי שהודפסו תוך כדי בצוע הנסויים השונים. תוצאות אלה נרשמו עבור שני טקסטים, ביחד עם המובנות ומספרי הנסוי בהם הם התקבלו. התוצאות תוארו גם בגרפים. בציור מס' C.2 מתוארת התלוח בין המובנות לשגיאה רבועית ממוצעת. בציור מס' C.3 מתוארת התלוח בין המובנות לקרוסקורלציה. בולטת מהן התרכזות התוצאות, עבור טקסט נחון, באזור מצומצם. עם זאת אין התאמה בין המובנות לכל אחד משני הנסויים.

טבלה C.1 אחוז הקווים הנבחרים בחחומי תדרים שונים

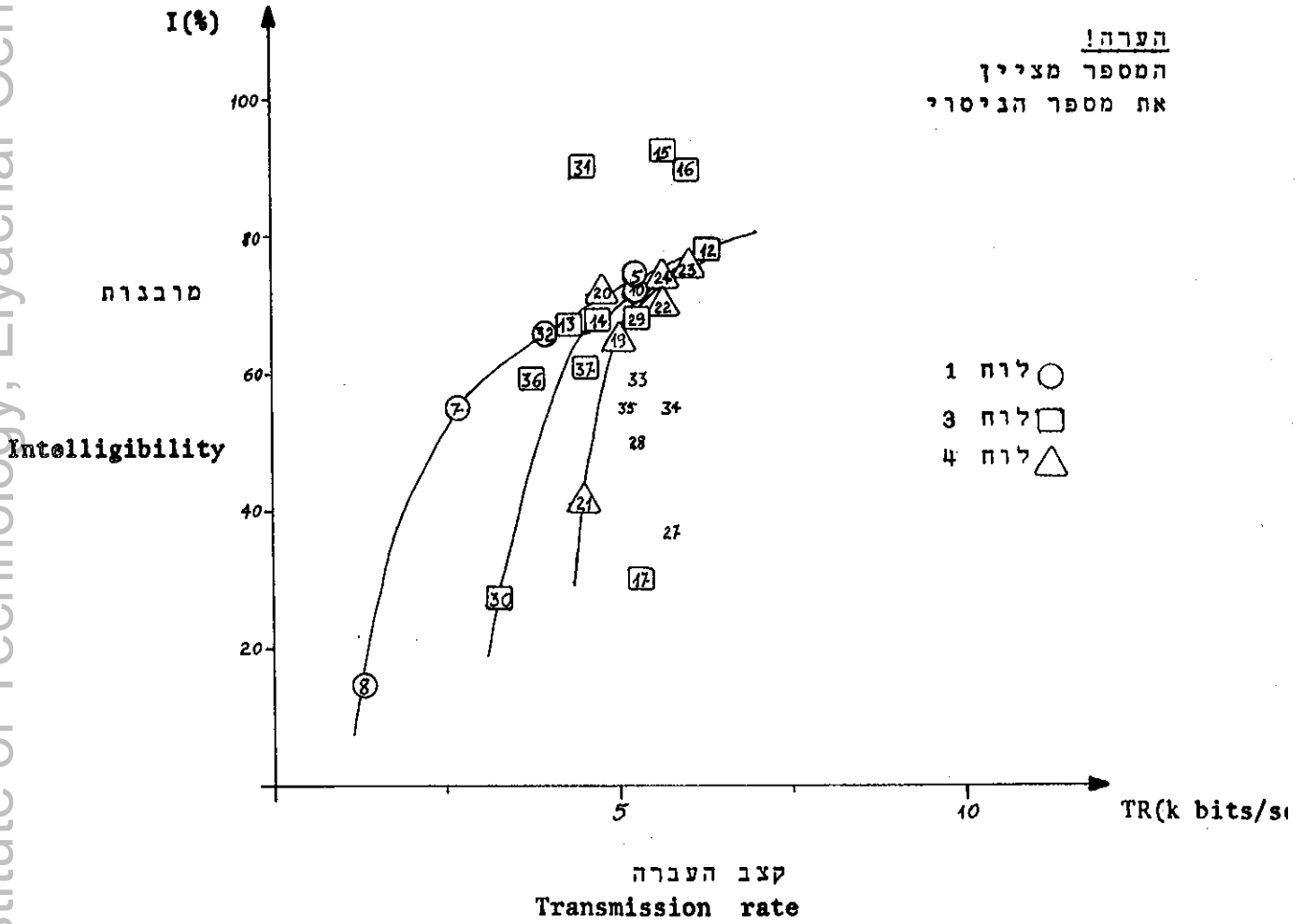
Table C.1. Percentage of chosen lines in different frequency bands

נסוי מס' Test No.	אחוז הקווים הנבחרים בחחומי תדרים שונים Percentage of chosen lines in different frequency bands		
	P(0) 300 - 1000Hz	P(1) 1000 - 2000Hz	P(2) 3000 - 4000Hz
13	62.7	28.4	8.8
15	67.1	25.5	7.4
16	71.4	23	5.6
17	60	30	10
19	82.2	13.7	4.1
20	89.2	8.7	2.1
22	78.2	16.8	5.0
23	82.7	14.2	3.1
24	67.4	25.4	7.2
28	55.9	37.4	6.7
29	55.3	33.7	11
32	60	31.8	8.2
33	51.2	32.4	16.4
ס"ה Total	883.3	321.0	95.6
ממוצע Average	68	24.7	7.3

טבלה C.2 קשר בין מובנות לקצב העברה

Table C.2. Relation of intelligibility to transmission rate

מובנות I(%)	קצב העברה TR(bits/sec)	נסוי מס' Test No.	מובנות I(%)	קצב העברה TR(bits/sec)	נסוי מס' Test No.
96	37800	4	58.9	5250	33
85.3	21000	6	73.9	5250	5
79	10500	3	54.7	5050	35
77.5	6250	12	64.7	5000	19
75.0	5937	23	67.1	4750	14
89.4	5937	16	71.7	4750	20
54.3	5750	34	41.1	4500	21
36.3	5750	27	89.2	4500	31
69.3	5625	22	60.2	4500	37
73.2	5625	24	66.9	4250	13
92.1	5625	15	65.2	3937	32
29.4	5328	17	58.7	3750	36
72	5250	10	27.0	3250	30
49.6	5250	28	54.2	2625	7
67.4	5250	29	14.6	1312	8



ציור מס C.1. תלות המרבנות בקצב ההעברה

Fig.C.1 Intelligibility as a function of transmission rate

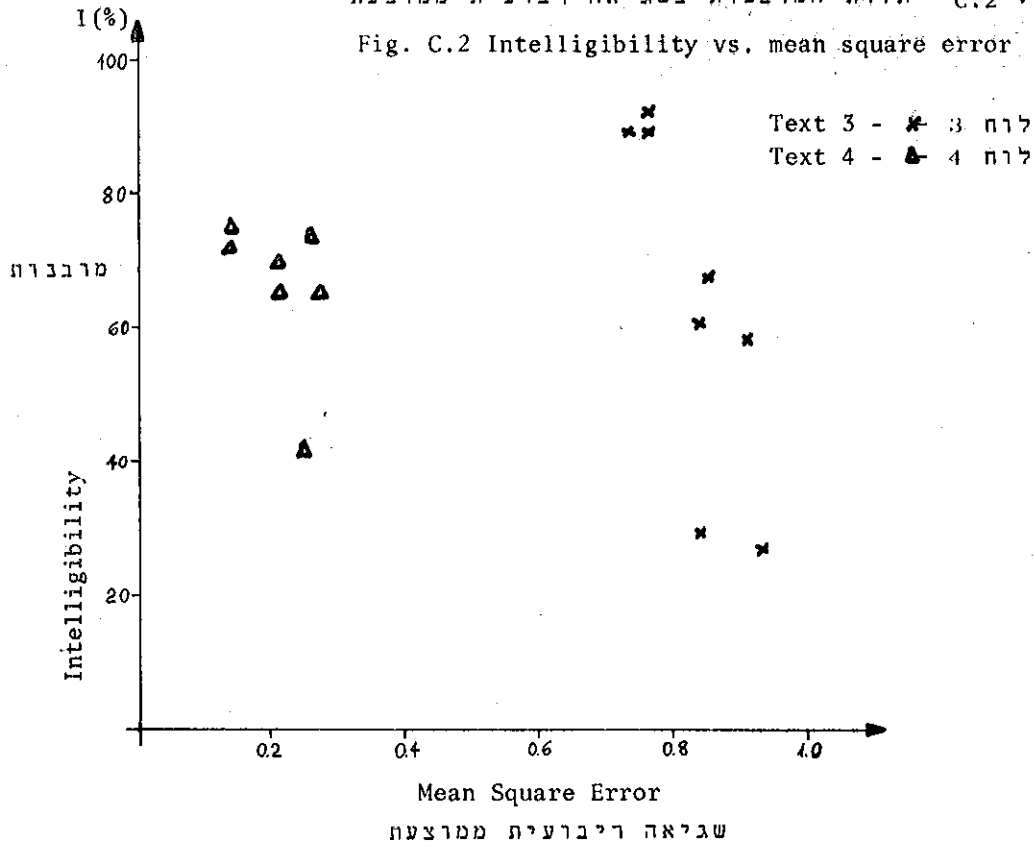
טבלה C.3 קשר בין קרוסקורלציה ושגיאה רבועית ממוצעת למובנות

Table C.3. The relation between crosscorrelation and mean square error to intelligibility

	נסוי מס' Test No.	מובנות Intelligibility I(%)	שגיאה רבועית ממוצעת Mean square error	קרוסקורלציה Crosscorrelation
לוח 3 Text 3	15	92.1	0.76	0.49
	16	89.4	0.73	0.53
	30	89.2	0.76	0.49
	29	67.4	0.85	0.39
	37	60.2	0.84	0.41
	36	58.7	0.91	0.35
	17	29.4	0.84	0.4
	31	27.0	0.93	0.34
לוח 4 Text 4	23	75.0	0.14	0.93
	24	73.2	0.26	0.86
	20	71.7	0.14	0.93
	22	69.3	0.21	0.89
	19	64.7	0.21	0.89
	21	41.1	0.25	0.87

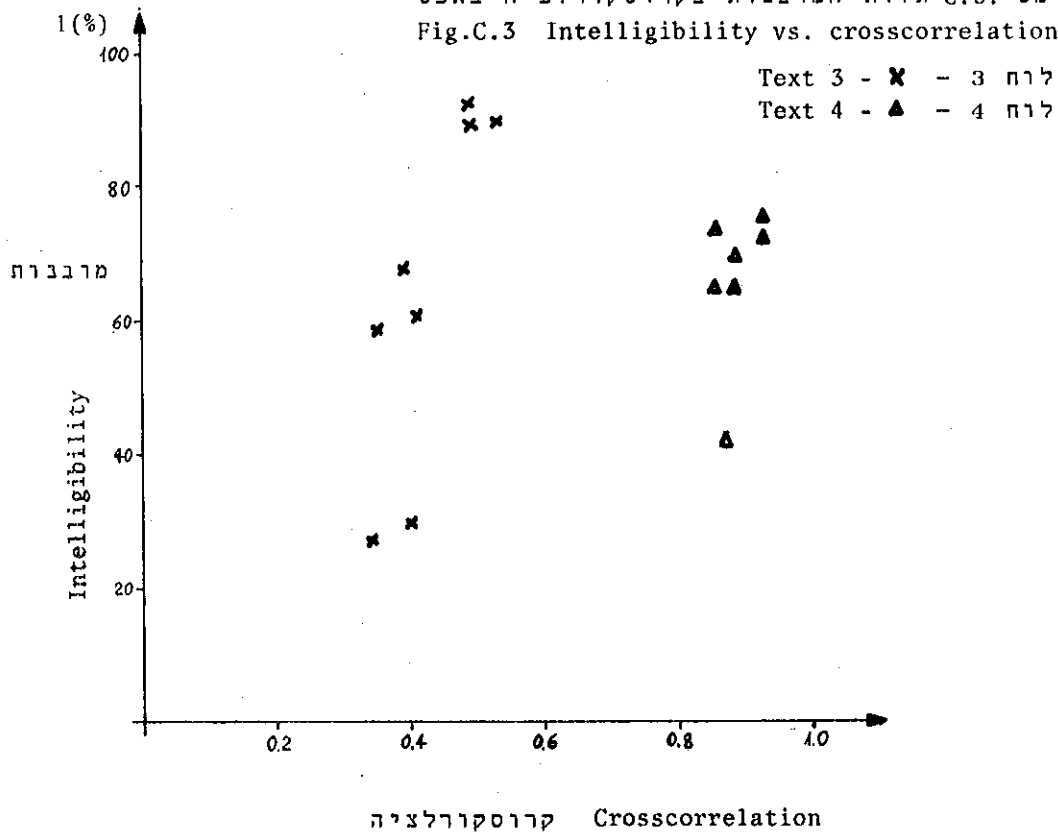
ציור מס. C.2 תלות המובנות בשגיאה רבועית ממוצעת

Fig. C.2 Intelligibility vs. mean square error



ציור מס. C.3 תלות המובנות בקרוסקורלציה באפס

Fig.C.3 Intelligibility vs. crosscorrelation at zero



נספח ד' - תכניות מחשב

1. מבוא

התכניות שנוצלו בעת הדמוי היו:

- א. תכנית דגימה והמרה אנלוגית לספרתית.
- ב. תכנית ראשית לבצוע העבוד.
- ג. תכנית העברה מסרט מגנטי לדיסק מגנטי.
- ד. תכנית המרה ספרתית לאנלוגית.

כל התכניות נכתבו בפורטרן IV. כל התכניות קוראות לשגרות (subroutines) הכתובות בשפת אסמבלר. התכניות הורצו במחשב IBM 360/44 אשר היה ברשות המכון למחקרי נפט וגיאופיסיקה, חולון.

בציור מס' 1. D מופיעה תכנית זרימה עקרונית עבור כל ארבע התכניות.

2. תכנית דגימה והמרה אנלוגית לספרתית

לבצוע קלט אנלוגי חובר המערך שבציור מס' 6 אל המחשב. לאחר הטענת התכנית נשמע צלצול פעמון אשר סימן שהמחשב מחכה כשהוא מוכן לבצע המרה. ההמרה החלה עם נתינת interrupt למחשב, ע"י המפעיל, בעזרת לחצן על מלואת הבקרה של המחשב. קצב הדגימה נקבע ע"י דפקי שעות שנמצא ב-interface של המחשב. עם סיום פעולת ההמרה נשמע הפעמון שנית כדי לאפשר הפסקת פעולת קלט הסרט האנלוגי. כמות הדגימות שנלקחו נקבעה ע"י כרטיס בקרה שכלל את המספר הרצוי.

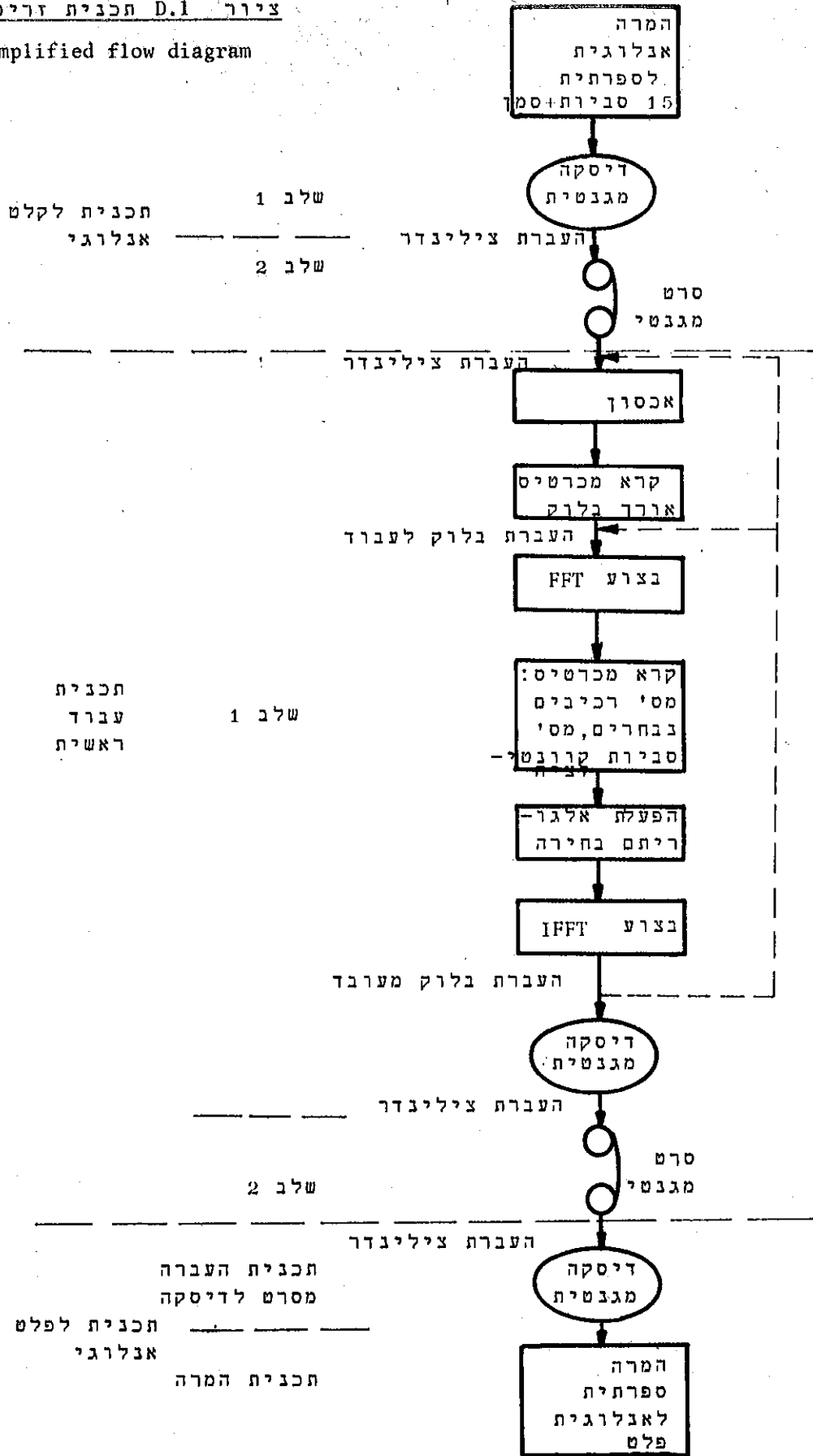
התכנית כוללת שני שלבים. בשלב ראשון בוצעה הדגימה, המרה אנלוגית לספרתית ואכסון הדגימות על דיסקה מגנטית. בשלב שני הועברו הדגימות מדיסקה מגנטית לסרט מגנטי. בהעברה זו הועבר כל פעם צילינדר שלם של הדיסקה (18000 מילים).

3. תכנית ראשית לבצוע העבוד

תכנית זו בצעה את העבוד שהוצע בסעיף 5.3.1. התכנית קראה שני כרטיסי בקרה אשר קבעו את מצב הפרמטרים של העבוד. תחילה הועבר תוכן של צילינדר מסרט מגנטי לזכרון של המחשב. לאחר מכן נקרא הכרטיס המחאר את אורך הבלוק N ואז נעשה עבוד של בלוק באורך זה. תחילה בוצעה התמרת פוריה לאחר מכן נקרא כרטיס המגדיר את n ואח z ואז הופעל

ציור D.1 תכנית זרימה עקרונית

Fig.D.1 Simplified flow diagram



בהחאם אלגוריתם הבחירה והקוונטיזציה. אלגוריתם זה אולגן כשגרה בשם 9 PREP. עם סיום הבחירה בוצעה התמרה הפוכה ותוכן הבלוק המקובד בדרך זו אוכסן על דסקה מגנטית. פעולה זו של העבוד חזרה על כל הבלוקים שאוכסנו בזכרון המחשב ועם סיומה הועבר תוכן צילינדר נוסף מן הסרט המגנטי לזכרון המחשב. כך נמשך העבוד עד סיום העבוד של כמות רצויה של דגימות. לאחר מכן עברה התכנית לשלב שני ואז הועבר תוכן הדיסקה לסרט מגנטי שני.

תוך כדי העבוד בוצעו הדפסות שונות וכן הדפסות בצורת גרפים. האחרונות בוצעו תוך ניצול שגרת Plot.

4. פלט אנלוגי

לבצוע פלט אנלוגי של הדגימות המעובדות המאוכסנות בסרט מגנטי היה צורך להשתמש בשתי תכניות.

התכנית הראשונה העבירה את תוכן הסרט המגנטי לדסקה מגנטית, כאשר בכל פעם מולא צילינדר שלם. לאחר סיום ההעברה לדיסקה הוטענה תכנית שניה אשר בצעה את ההמרה מספרתי לאנלוגי. לבצוע הפלט האנלוגי חובר המערך שבציוור מס' 7, אל המחשב. לאחר הטענת התכנית נשמע צלצול פעמון אשר סימן למפעיל שהמחשב מחכה ומוכן לבצוע ההמרה. ההמרה החלה עם נתינת interrupt למחשב, ע"י המפעיל, בעזרת לחצן על מלואת הבקרה של המחשב. עם סיום ההמרה נשמע הפעמון שנית.

```
// JOB ,EJM
IA55I 002043
//SYS002 ACCESS INP,192=
//SYS004 ACCESS OUT1,180=
//SYS005 ACCESS OUT,180=
//SYS008 ACCESS WORK2,291='333333'
IA89I M 291 333333
// EXEC FORTRAN
```

תכנית דגימה
וקמרה אנלוגית
לספרות

FORTRAN IV MODEL 44 PS VERSION 5, LEVEL 0 DATE 11/11/61

```
0001            INTEGER*2 TABLE1(18000),TABLE2(18000)
              *,IBUF(36000)
0002            DIMENSION KHEAD(2)
0003            EQUIVALENCE (TABLE1,IBUF),(TABLE2,IBUF(18001))
0004            KHEAD(1)=0
0005            55 CONTINUE
0006            READ (1,101,END=202) N
0007            CALL IOOP(7,2,TABLE1,3600,0,3600)
0008            CALL DELAY(7,IY,IYY)
0009            KHEAD(1)=KHEAD(1)+1
0010            KHEAD(2)=N
0011            CALL IOOP(3,1,KHEAD,8)
0012            CALL DELAY(3,IP,IPP)
0013            PRINT 77,KHEAD(1),KHEAD(2)
0014            77 FORMAT (20X,'TEST-NO.',I3,4X,'SAMPLES=',I7)
0015            N1=N
0016            101 FORMAT (I10)
0017            MDD=MOD(N,18000)
0018            IF (MDD.EQ.0) GO TO 2
0019            N1=(N/18000+1)*18000
0020            2 CONTINUE
              C**** N IS THE NUMBER OF READ-OPERATIONS
              C**** NCCW IS THE NUMBER OF 500-CCW-TABLES
0021            NCCW=N1/500
0022            CALL BELL
0023            CALL AROUT(NCCW,TABLE1,TABLE2)
0024            CALL BELL
0025            NB=36000
0026            57 CONTINUE
0027            CALL IOOP(7,2,IBUF,36000,I,3600)
0028            CALL DELAY(7,IZ,IZZ)
0029            CALL IOOP(4,1,IBUF,NB)
0030            CALL DELAY(4,IY,IYY)
0031            N=N-18000
0032            I=I+10
0033            IF (N.GE.18000) GO TO 57
0034            IF(N.LE.0) GO TO 55
0035            NB=2*N
0036            GO TO 57
0037            202 STOP
0038            END
```

// JOB ,EJM
IA55I 014953
// EXEC FORTRAN

מכנת ראשית
לצורך חשבון

```

0001      INTEGER*2  INP(18700)
0002      INTEGER*2  TEMP(1024)/1024*0/
0003      DIMENSION  KEEP(1024),TABLE(1024),X(1024),Y(1024),STOR(
0004 1000 CONTINUE
0005      READ (1,2,END=3) JBF
C**** JBF IS PROCESSED-BLOCK-LENGTH (POWER OF 2)
0006      2 FORMAT (I4)
0007      PRINT 517,JBF
0008      517 FORMAT ('1',20X,'BLOCK-LENGTH IS',I5,' HALF-WORDS')
0009      JBF1=JBF
0010      IPOW=0
0011      4 JBF1=JBF1/2
0012      IPOW=IPOW+1
0013      IF (JBF1.GT.1) GO TO 4
C****JBF=2**IPOW
0014      LIM=1024/JBF
0015      ITRACK=0
0016      ISOF=18000
0017      ISTART=1
0018      IBLOC=0
0019      CALL IOOP(1,2,INP,3600)
0020      CALL DELAY(1,IX,IXX)
0021      1 CALL IOOP(1,2,INP(ISTART),36000)
0022      CALL DELAY(1,IX,IXX)
0023      301 FORMAT (18(1X,I6))
0024      IF (IX.LT.0) GO TO 17
0025      IEND=17
0026      IF (ISOF.GE.18432) IEND=18
0027      TRMSE=0.
0028      TCORR=0.
0029      TSUM1=0.
0030      TSUM2=0.
0031      DO 10 I=1,IEND
0032      JIN=1+1024*(I-1)
0033      JSTART=JIN
0034      DO 20 ILIM=1,LIM
0035      JJJ=ILIM
0036      IBLOC=IBLOC+1
0037      CALL SCALH(INP,JIN,INP,JIN,JBF,-3)
0038      DO 21 J=1,JBF
0039      KK=JIN+J-1
0040      21 STOR(1,J)= INP(KK)/60.
0041      CALL FFTIP(INP(JIN),INP(JIN),X,Y,IPOWER,KEEP,
*TABLE,-1.0,IFLAG,ISUBT,F,G,NC,JBF)
0042      CALL FREDOM(X,Y,IFLAG,ISUBT,F,G,NC,JBF)
0043      IF (IFLAG.LT.0) STOP 12
0044      CALL FFTIP(INP(JIN),INP(JIN),X,Y,IPOWER,KEEP,
*TABLE, 1.0,IFLAG,ISUBT,F,G,NC,JBF)
C      COMPUTES MEAN SQUARE ERROR AND CORRELATION FOR PROCESSI
C      UNPROCESSED BLOCK AND PLOT THEM ON ONE GRAPH IN TIME DC
0045      RMSE=0.
0046      SUM1=0.
0047      SUM2=0.
0048      CORR=0.
0049      DO 22 J=1,JBF
0050      KK=JIN+J-1
0051      STOR(2,J)= INP(KK)/60.
0052      RMSE=RMSE+(STOR(1,J)-STOR(2,J))**2
0053      CORR=CORR+STOR(1,J)*STOR(2,J)

```



```

0054      SUM1=SUM1+STOR(1,J)**2
0055      SUM2=SUM2+STOR(2,J)**2
0056      22 CONTINUE
0057      LL=13*512/JBF
0058      LH=17*512/JBF
0059      IF( IBLOC.NE.LL.AND. IBLOC.NE.LH) GO TO 24
0060      SKAL=SQRT(SUM1*SUM2)
0061      PRMSE=RMSE/SUM1
0062      PCORR=CORR/SKAL
0063      PRINT 23, PRMSE, PCORR
0064      23 FORMAT (1H ,26H BLOC MEAN SQUARE ERROR IS,F15.6, /
      *1H ,20H BLOC CORRELATION IS,F15.6)
0065      CALL PLOT (STOR,2,JBF,50)
0066      24 TRMSE=TRMSE+RMSE
0067      TCORR=TCORR+CORR
0068      TSUM1=TSUM1+SUM1
0069      TSUM2=TSUM2+SUM2
0070      IF( IBLOC.NE.(20480/JBF)) GO TO 26
0071      SKAL=SQRT(TSUM1*TSUM2)
0072      TRMSE=TRMSE/TSUM1
0073      TCORR=TCORR/SKAL
0074      PRINT 25, TRMSE, TCORR
0075      25 FORMAT(1H ,32H TOTAL ROOT MEAN SQUARE ERROR IS,F15.6, /
      *1H ,21H TOTAL CORRELATION IS,F15.6)
0076      26 CONTINUE
0077      JIN=JIN+JBF
0078      IF (IFLAG.NE.0) GO TO 33
0079      20 CONTINUE
0080      33 CONTINUE
0081      IF (JJJ.EQ.LIM) GO TO 44
0082      J24=JSTART+1023
0083      DO 55 LLL=JIN,J24
0084      55 INP(LLL)=0
0085      44 CONTINUE
0086      ITRACK=ITRACK+1
0087      CALL IOOP(7,1,INP(JSTART),2048,ITRACK,2048)
0088      CALL DELAY(7,IY,IYY)
0089      IF (IFLAG.EQ.0) GO TO 10
0090      DO 13 M=1,10
0091      ITRACK=ITRACK+1
0092      CALL IOOP(7,1,TEMP,2048,ITRACK,2048)
0093      CALL DELAY(7,IY,IYY)
0094      13 CONTINUE
0095      CALL IOOP(1,7)
0096      CALL DELAY(1,IX,IXX)
0097      NBT=2048*ITRACK
0098      PRINT 15, ITRACK, NBT
0099      15 FORMAT (' TRACKS USED*I10,I10)
0100      NHW=0
0101      NST=1
0102      ISW1=0
0103      DO 30 K=1, ITRACK
0104      CALL IOOP(7,2,INP(NST),2048,K,2048)
0105      CALL DELAY(7,IZ,IZZ)
0106      NHW=NHW+1024
0107      NST=NST+1024
0108      IF (NHW.LT.18000) GO TO 30
0109      DO 40 L=1,18000
0110      40 INP(L)=INP(L)+16383

```

```
0111      CALL IOOP(2,1,INP,36000)
0112      CALL DELAY(2,IW,IWW)
0113      NHW=NHW-18000
0114      NST=NHW+1
0115      CALL MOVEHW(INP(18001),INP(1),NHW)
0116  30 CONTINUE
0117      DO 190 IJ=1,NHW
0118  190   INP(IJ)=INP(IJ)+16383
0119      NHWP1=NHW+1
0120      DO 200 IJ=NHWP1,18000
0121  200   INP(IJ)=16383
0122      CALL IOOP(2,1,INP,36000)
0123      CALL DELAY(2,IW,IWW)
0124      GO TO 1000
0125  10 CONTINUE
0126      JSTART=JSTART+1024
0127      NHW=ISOF+1-JSTART
0128      CALL MOVEHW(INP(JSTART),INP(1),NHW)
0129      ISOF=18000+NHW
0130      ISTART=NHW+1
0131      GO TO 1
C**** END OF FFT
0132  17 CONTINUE
0133      PRINT 100
0134  100  FORMAT(' E.O.F. FOUND ON INPUT ')
0135      STOP 456
0136      3 PRINT 112
0137  112  FORMAT ('OEND OF PROCESSING')
0138      STOP 999
0139      END
```

```

001      SUBROUTINE FFTIP(IN,IY,X,Y,N,KEEP,TABLE,SIGNI,
002      *IFLAG,ISUBT,F,G,NC,JBF)
      DIMENSION IN(1),IY(1),X(1),Y(1),KEEP(1),TABLE(1)
C      IN-HALFWORD INPUT ARRAY(INTEGER*2) FOR DIRECT F.T.
C      IY-HALFWORD OUTPUT ARRAY(INTEGER*2) FOR INVERSE F.T.
C      X -REAL PART OF SERIES (OUTPUT FOR FOURIER TRANS FORM) (R)
C      INPUT FOR INVERSE FOURIER TRANSFORM
C      Y -IMAGINARY PART OF SERIES (OUTPUT FOR DIRECT F.T.)
C      INPUT FOR INVERSE F.T.) REAL
C      N -THE POWER OF 2 NUCH THAR 2**N IS THE LENGTH OF TIME SEP
C      KEEP- WORK AREA (INTEGER*4
C      TABLE - WORK AREA (REAL*4
C      SIGNI=-1.0 FOR DIRECT F.T
C      1.0 FOR INVERSE F.R.
C      IN ,IY,X,Y,KEEP,TABLE-MUST BE DIMENSIONED AT LEAST 2**N
C      -----
003      DATA NP/0/
004      IF(INP.EQ.N) GO TO 1
005      NP=N
006      LX=2**N
007      CALL FASTAB(N,KEEP,TABLE)
008      1 IF(SIGNI.GT.0) GO TO 2
009      CALL H2FL('CPU',IN,LX,X)
010      2 CALL FFT(X,Y,N,LX,SIGNI,KEEP,TABLE,IER)
011      IF(IER.NE.0) STOP 12345
012      IF(SIGNI.LT.0) RETURN
013      CALL FL2H ('CPU',X,LX,1.,1.,IY,32767)
014      RETURN
015      E N D

```

FORTRAN IV

MODEL 44 PS

VERSION 5, LEVEL 0 DATE

```

0001      SUBROUTINE FREDDOM(XRE,XIM,IFLAG,ISUBT,F,G,NC,JBF)
0002      DIMENSION XRE(512),XIM(512)
0003      DATA IBLOC/0/
0004      IFLAG=0
0005      IBLOC=IBLOC+1
0006      IF(MOD(IBLOC,10).EQ.0) PRINT 55,IBLOC
0007 55  FORMAT (1H ,I10)
0008      IF(IBLOC.NE.1) GO TO 20
0009      READ(1,50,END=40) ISUBF,D,E,ISUBT,F,G,NC,JCH
0010      PRINT 51,ISUBF,D,E,ISUBT,F,G,NC,JCH
0011 50  FORMAT(I10,2E10.0,I10,2E10.0,2I10)
0012 51  FORMAT (1H0,I10,2F10.0,I10,2F10.0,2I10)
0013      BL=76800/JBF
0014      FACT=2.**(1./BL)
0015 20  IF (JCH.EQ.1) D=D*FACT
0016      GO TO (1,2,3,4,5,6,7,8,9,10), ISUBF
0017      1 CALL FREP1(XRE,XIM,D,E,JBF)
0018      GO TO 10
0019      2 CALL FREP2(XRE,XIM,D,E,JBF)
0020      GO TO 10
0021      3 CALL FREP3(XRE,XIM,D,E,JBF)
0022      GO TO 10
0023      4 CALL FREP4(XRE,XIM,D,E,JBF)
0024      GO TO 10
0025      5 CALL FREP5(XRE,XIM,D,E,JBF)
0026      GO TO 10
0027      6 CALL FREP6(XRE,XIM,D,E,JBF)
0028      GO TO 10
0029      7 CALL FREP7(XRE,XIM,D,E,JBF)
0030      GO TO 10
0031      8 CALL FREP8(XRE,XIM,D,E,JBF)
0032      GO TO 10
0033      9 CALL FREP9(XRE,XIM,D,E,JBF)
0034 10  IF (IBLOC.NE.(20480 /JBF)) GO TO 30
0035      IFLAG=1
0036      IBLOC=0
0037 30  RETURN
0038 40  IFLAG=-1
0039      RETURN
0040      END

```

FORTRAN IV

MODEL 44 PS

VERSION 5, LEVEL 0

```

0001      SUBROUTINE FREP9 (XRE,XIM,D,E,JBF)
0002      DIMENSION XRE(1),XIM(1),ST(40),XX(513),MEM(513),R(3),SX
0003      DATA IBLOC /0/,R/3*0./
0004      WRITE (3,1)
0005      1 FORMAT (1H0)
0006      IBLOC=IBLOC+1
0007      ID=D
0008      IE=E
0009      FM=18.*ALOG10(2.)/(2.**IE)
0010      IBF=(JBF/2)+1
0011      IF (IBLOC.NE.1) GO TO 3
0012      DO 4 K=1,IBF
0013      4 MEM(K)=0
0014      3 CONTINUE
0015      IL=JBF*3/80
0016      IH=JBF*3/8
0017      XRE(1)=0.
0018      XIM(1)=0.
0019      DO 5 K=2,IL
0020      J=JBF+2-K
0021      XRE(K)=0.
0022      XRE(J)=0.
0023      SXX(1,K)=0.
0024      SXX(2,K)=0.
0025      XIM(K)=0.
0026      5 XIM(J)=0.
0027      DO 6 K=IH,IBF
0028      J=JBF+2-K
0029      XRE(K)=0.
0030      XRE(J)=0.
0031      SXX(1,K)=0.
0032      SXX(2,K)=0.
0033      XIM(K)=0.
0034      6 XIM(J)=0.
0035      DO 10 K=1,ID
0036      10 ST(K)=0.
0037      ISL=1
0038      PLIN=0.
0039      JSET=0
0040      IF(ID.EQ.(JBF/2))RETURN
C          PRINT UNPROCES BLOC NO. LL AND LH
0041      IF(IBLOC.NE.1) GO TO 707
0042      READ 11,IBL,IDOT,IPLLOT,IPRUN
0043      11 FORMAT (2A1,2I1)
0044      707 CONTINUE
0045      IF(IPRUN.EQ.0.AND.IPLOT.EQ.0)GOTO13
0046      LL=16*512/JBF
0047      LH=17*512/JBF
0048      IF(IBLOC.NE.LL.AND.IBLOC.NE.LH)GO TO 13
0049      DO 12 K=IL,IH
0050      J=K-IL+1
0051      XX(K)=SQRT(XRE(K)**2+XIM(K)**2)
0052      IF (IPRUN.EQ.0)GO TO 112
0053      PRINT 37,K,XX(K),XRE(K),XIM(K)
0054      112 SXX(1,K)=XX(K)/2200.
0055      12 CONTINUE
0056      IF(IPLLOT.EQ.0) GO TO 13
0057      KSO=JBF/2
0058      13 CONTINUE

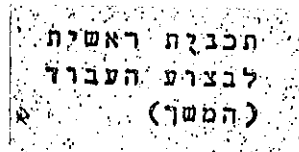
```

FORTRAN IV MODEL 44 PS VERSION 5, LEVEL 0

```

0059      DO 30 K=IL,IH
0060      XX(K)=XRE(K)**2+XIM(K)**2
0061      JSET=JSET+1
0062      IB=0
0063      IF (XX(K).LT.PLIN) IB=1
0064      IF (ISL.EQ.1.AND.IB.EQ.1) GO TO 15
0065      XX(K-1)=0.
0066      GO TO 25
15 JSET=0
0068      DO 20 J=1,ID
0069      I=ID-J+1
0070      IF (PLIN.LE.ST(I)) GO TO 25
0071      ST(I+1)=ST(I)
0072      20 ST(I)=PLIN
0073      25 PLIN=XX(K)
0074      ISL=1
0075      IF (IB.EQ.1) ISL=0
0076      30 CONTINUE
0077      P=ST(ID)
0078      DO 40 K=IL,IH
0079      M=0
0080      J=JBF-K+2
0081      IF (XX(K).GE.P.AND.XX(K).GT.0.) GO TO 35
0082      32 XRE(K)=0.
0083      XRE(J)=0.
0084      XIM(K)=0.
0085      XIM(J)=0.
0086      SXX(2,K)=0.
0087      GO TO 40
0088      35 F=ABS(XRE(K))+1.
0089      F=ALOG10(F)/FM
0090      N=F
0091      F=N
0092      F=10**(F*FM)
0093      XRE(K)=SIGN(F,XRE(K))
0094      XRE(J)=XRE(K)
0095      F=ABS(XIM(K))+1.
0096      F=ALOG10(F)/FM
0097      N=F
0098      F=N
0099      F=10**(F*FM)
0100      XIM(K)=SIGN(F,XIM(K))
0101      XIM(J)=-XIM(K)
0102      MEM(K)=MEM(K)+1
0103      36 CONTINUE
0104      XX(K)=SQRT(XX(K))
0105      LB=LL-4
0106      LS=LH+4
0107      IF (IBLOC.GT.LB.AND.IBLOC.LT.LS) WRITE(3,37)K,XX(K),XRE(K)
0108      37 FORMAT (1H0,I5,3(5X,F10.0))
0109      SXX(2,K)=XX(K)/2200.
0110      IF (M.EQ.0) GO TO 40
0111      XX(K)=0.
0112      40 CONTINUE
C      PLOT OF PROC BLOC NO.LL AND LH
0113      IF (IPLOT.EQ.0) GO TO 41
0114      IF (IBLOC.EQ.LL.OR.IBLOC.EQ.LH) CALL PLOT(SXX,2,KSD,100)
0115      41 CONTINUE
0116      IF (IBLOC.NE.(153600/JBF)) RETURN

```



FORTRAN IV MODEL 44 PS VERSION 5, LEVEL 0

```
C            STATISTIC OF CHOSEN LINES
0117            L1=JBF/8
0118            L2=L1+1
0119            L3=JBF/4
0120            L4=L3+1
0121            DO 46 I=L1,L1
0122            46 MEM(1)=MEM(1)+MEM(I)
0123            R(1)=MEM(1)*100./(ID*IBLOC)
0124            DO 47 I=L2,L3
0125            47 MEM(2)=MEM(2)+MEM(I)
0126            R(2)=MEM(2)*100./(ID*IBLOC)
0127            DO48 I=L4,IH
0128            48 MEM(3)=MEM(3)+MEM(I)
0129            R(3)=MEM(3)*100./(ID*IBLOC)
0130            PRINT 50,(I,R(I),I=1,3)
0131            50 FORMAT(1H , 'PERCENT OF LINES CHOSEN IN REG',I1,F10.2)
0132            IBLOC=0
0133            RETURN
0134            END
```

FORTRAN IV

MODEL 44 PS

VERSION 5, LEVEL 0

```

0001      SUBROUTINE PLOT (Y,M,NF,NS)
C        SUBROUTINE FOR PLOTTING 5 X 100 INPUT ARRAY (FORTRAN
C          Y=TWO DIMENSIONAL ARRAY OF VARIABLES Y(I,J) GIVING
C            VALUE OF THE I-TH FUNCTION WHICH IS TO BE PLOTTED
C          M=NUMBER OF QUANTITIES TO BE PLOT
C          NF=NUMBER OF LINES TO BE PLOT
C          NS=THE MAXIMUM VALUE FOR THE ORDINATE SCALE
C          -----
0002      DIMENSION Y(2,100),LINE(101),L(11),JL(5)
0003      DATA          JL/1HA,1HB,1HC,1HD,1HE/,JN,JP,JI,JBLA
0004      DO 99 I=1,101
0005      LINE(I)=JBLANK
0006      99 CONTINUE
0007      N=0
C        PRINT ORDINATE SCALE
0008      DO 101 I=1,11
0009      L(I)=10*I-110+NS
0010      101 CONTINUE
0011      PRINT 105,(L(I),I=1,11)
0012      105 FORMAT (3X,11(I4,6X),6HY(1,I))
0013      GO TO 115
0014      110 IF (N/10-(N-1)/10) 125,125,115
C        CONSTRUCT ORDINATE GRAPH LINE
0015      115 ND=0
0016      DO 120 I=1,10
0017      ND=ND+1
0018      LINE(ND)=JP
0019      DO 120 J=1,9
0020      ND=ND+1
0021      120 LINE(ND)=JN
0022      LINE(101)=JP
0023      IF (N) 135,121,135
0024      121 PRINT 170,N,LINE
0025      GO TO 185
C        CONSTRUCT 1 LINE OF ABSCISSA GRAPH LINES
0026      125 DO 130 I=1,101,10
0027      LINE(I)=JI
0028      130 CONTINUE
C        CHANGE NUMERICAL DATA TO LETTERS
0029      135 DO 160 I=1,M
0030      XNS=NS
0031      JA=Y(I,N)+101.49999-XNS
0032      IF (JA-101) 140,155,145
0033      140 IF (JA) 150,150,155
0034      145 LINE(101)=JZ
0035      GO TO 160
0036      150 LINE(1)=JZ
0037      GO TO 160
0038      155 LINE(JA)=JL(I)
0039      160 CONTINUE
C        PRINT LINE OF DATA
0040      IF (N/10-(N-1)/10) 175,175,165
0041      165 PRINT 170,N,LINE,Y(1,N)
0042      170 FORMAT (1X,I4,101A1,1X,1PE12.5)
0043      GO TO 185
0044      175 PRINT 180,LINE,Y(1,N)
0045      180 FORMAT (5X,101A1,1X,F12.5)
C        SET LINE VARIABLES TO ZERO

```


FORTRAN IV MODEL 44 PS VERSION 5, LEVEL 0

```
0046            185 DO 190 I=1,101  
0047                LINE(I)=JBLANK  
0048            190 CONTINUE  
0049            195 N=N+1  
0050                IF (N-NF) 110,110,200  
0051            200 RETURN  
0052                END
```

// JOB ,TPTDSK
IA55I 000031
//SYS002 ACCESS INP,181=
// EXEC FORTRAN

תכנית העברה
מסרט מגנטי
לדיסק מגנטי

FORTRAN IV

MODEL 44 PS

VERSION 5, LEVEL 0

```
0001      INTEGER*2  IBUF(18000)
0002      IUI=3*16383
0003      ITRACK=1
0004      NSAMP=1000
0005      1 CALL  ICOP(1,2,IBUF,36000,-22)
0006      CALL  DELAY(1,IX,IXX)
0007      NB=36000-IXX
0008      IF (NB.LT.1000) GO TO 1
0009      IF (IX.NE.0) GO TO 2
0010      DO 101 JJ=1,18000
0011      101  IBUF{JJ)=(IBUF{JJ)-16383)*2+16383
0012      CALL  ICOP(7,1,IBUF,36000,ITRACK,3600)
0013      CALL  DELAY(7,IY,IYY)
0014      ITRACK=ITRACK+10
0015      NSAMP=NSAMP+18000
0016      GO TO 1
0017      2 PUNCH 3,NSAMP
0018      3 FORMAT (I10)
0019      END
```

```

// JOB ,MMM
IA55I 002735
//SYS008 ACCESS WORK2,291='333333'
IA89I M 291 333333
//SYS002 ACCESS AQ,195=
// EXEC ANOUTPT
STOP 0

```

תכנית המרה
ספרתית לאנלוגית

1. A.D. Arkhipova and M.A. Sapozhkov, "The quality of Vocoder speech", Soviet Physics - Acoustics, vol. 16, no. 3, pp. 292 - 298, Jan. - Mar. 71.
2. E.O. Brigham and R.E. Morrow, "The fast Fourier transform", IEEE Spectrum, vol. 4 pp. 63 - 70, Dec. 67.
3. S.J. Campanella and G.S. Robinson, "A comparison of orthogonal transformations for digital speech processing", IEEE Trans. Communication Technology, vol. COM-19, no. 6, pp. 1045-1049, Dec. 71.
4. W.T. Cochran et al., "What is the fast Fourier transform", IEEE Trans. Audio and Electroacoustics, vol. AU-15, no. 2, pp. 45-55, June 67.
5. J.W. Cooley, P.A.W. Lewis and P.D. Welch, "The finite Fourier transform", IEEE Trans. Audio and Electroacoustics, vol. AU-17, no. 2, pp. 77-85, June 69.
6. W.R. Crowther and C.M. Rader, "Efficient coding of vocoder channel signals using linear transformation", Proc. IEEE, vol. 54, pp. 1594-1595, Nov. 1966.
7. G. Fant and K.N. Stevens, "Systems for speech compression", Fortschritte der Hochfrequenztechnik, vol. 5, pp. 229-262, 1960.
8. J.L. Flanagan, Speech Analysis, Synthesis and Perception. New York and Berlin: Springer-Verlag, 1965.
9. J.L. Flanagan, "Focal points in speech communication research", IEEE Trans. Communication Technology, vol. COM-19, no. 6, pp. 1006-1015, Dec. 71.
10. N.R. French and J.C. Steinberg, "Factors governing the intelligibility of speech sounds", J. Acoust. Soc. Am., vol. 19, no. 1, pp. 90-119, Jan, 1947.
11. W.M. Gentleman and G. Sande, "Fast Fourier transforms - for fun and profit", 1966 Fall Joint Computer Conference, AFIPS Proc., vol. 29, pp. 563-578.
12. B. Gold and C.M. Rader, Digital Processing of Signals. New York: McGraw-Hill, 1969.
13. B. Gold and C.M. Rader, "The channel vocoder", IEEE Trans. Audio and Electroacoustics, vol. AU-15, no. 4, pp. 148-161, Dec. 67.
14. W.D. Larkin and L.C. Stewart, "High quality channel vocoder", J. Acoust. Soc. Am., vol. 50, p. 107 (Abstract), 1971.
15. I. Lehiste and G.E. Peterson, "Linguistic considerations in the study of speech intelligibility", J. Acoust. Soc. Am., vol. 31, no. 3, pp. 280-286, Mar. 1959.
16. J.E. Luck, H.J. Manley and H.L. Shaffer, "Speech compression by component selection in the frequency domain", J. Acoust. Soc. Am., vol. 50, p. 107 (Abstract), 1971.

17. J. Max, "Quantizing for minimum distortion" IRE Trans. on Information Theory, vol. IT-6, no. 1, pp. 7-12, Mar. 60.
18. R.K. Otnes and L. Enochson, Digital Time Series Analysis, New York: John Wiley and Sons, 1972.
19. E. Peterson and F.S. Cooper, "Peakpicker: a bandwidth compression device", J. Acoust. Soc. Am., vol. 29, p. 777 (Abstract), 1957.
20. I. Pollack, H. Rubenstein and L. Decker, "Intelligibility of known and unknown message sets", J. Acoust. Soc. Am., vol. 31, no.3, pp. 273-279, Mar. 1959.
21. R.W. Schaffer, "A survey of digital speech processing techniques", IEEE Trans. Audio and Electroacoustics, vol. AU-20, no. 1, pp. 28-35, Mar. 72.
22. M.R. Schroeder, "Vocoders: Analysis and synthesis of speech", Proc. IEEE, vol.54, pp. 720-734, May 1966.
23. F.Y.Y. Shum, A.R. Elliott and W.O. Brown, "Speech processing with Walsh - Hadamard transforms", IEEE Trans. Audio and Electroacoustics, vol. AU-21, no. 3, pp. 174-179, June 73.

Speech Compression by means of
the Discrete Fourier Transform

Project Thesis

Submitted in partial fulfilment of the requirements
for the degree of Master of Science
in Electrical Engineering

By

Jakov Shulman

Submitted to the Senate of the Technion-Israel Institute of Technology

Tishri 5735

Haifa

September 1974

The work was carried out under the guidance of Dr. David Malach from the Dept. of Electrical Engineering of the Technion - Israel Institute of Technology.

I am very grateful for his constant support.

I would like to thank Mr. A. Gutman for his encouragement to utilize this subject for a Master's Thesis and Dr. M. Bitran for valuable discussions in the course of the work.

Many thanks go also to the Management of Eljim Ltd. for supporting the work.

Contents

	Page
1. Summary	1
2. List of symbols and abbreviations	3
3. Introduction	4
4. Compression of speech signals	5
4.1 The speech	5
4.2 Compression of speech signals using analysis - synthesis methods	11
4.3 Quality of reconstructed speech	14
5. Compression of speech signals based on Fourier Transform	17
5.1 Discrete Fourier Transform and its properties	17
5.2 Fast Fourier Transform	19
5.3 Using Fourier Transform for speech compression	19
5.4 Speech compression based on different orthogonal transforms	24
6. Description of the work and results	25
6.1 Computer simulation	25
6.2 Intelligibility tests	27
6.3 Intelligibility as a function of three basic parameters	29
6.4 Influence of Hann window	35
6.5 Transmission rate and intelligibility	37
6.6 Influence of different changes in the reduction method	39
6.7 Quality of the reconstructed speech	44
6.8 Intelligibility of synthesized speech, crosscorrelation and mean square error	44
7. Conclusions and recommendations	47
7.1 Summary of optimal conditions for local peaks method	47
7.2 Recommendations for continuation	48
8. Appendices	49
Appendix A. Details of tests and the results of intelligibility tests	49
Appendix B. Words used for tests	56
Appendix C. Collection of different results	57
C.1 Percentage of selected lines in three bands	
C.2 Transmission rate in all tests	
C.3 Crosscorrelation and mean square error	
Appendix D. Computer programs	63
9. References	81

Summary

Speech signal is used to transfer messages in speech communication links. The channel capacity required for the transmission of this signal is several orders higher than the rate at which information can be generated or processed by human beings. Reduction of the redundancy in the speech signal will reduce the required channel capacity.

Many analysis - synthesis methods are known for achieving this reduction. They are based either on the knowledge of the speech production mechanism and measurements of related parameters (excitation and filter functions) or on the shape of the voice spectrum and measurement of related parameters.

The aim of this work was to study the possibility of speech signal reduction using Discrete Fourier Transform (DFT) but without measurement of excitation function parameters. As the ear is very sensitive to small changes in the excitation function it is a great advantage to have a reduction method based on other parameters.

The suggested local maxima method is based on analysis of speech signal using DFT and compression by selection of the dominant components of this transform, while setting all the others to zero. The selection method is based on the shape of the discrete spectrum. The chosen transform components are those being the highest local maxima of the spectrum. The complex amplitude and spectral location for each of the selected components are transferred to the synthesizer, in digital form. Logarithmic quantization of the amplitude is used. For the synthesis of the speech signal inverse DFT is applied.

Computer simulation of the suggested method was performed using the fast Fourier transform algorithm. The intelligibility of the synthesized speech was measured and the influence of three basic parameters has been checked. These parameters are:

the number of selected components, the DFT block length (no. of samples), the resolution of the logarithmic quantization. The intelligibility was calculated from results of listening tests using 15 Hebrew words of two syllables each. 30 tests with 10 listeners in each test were performed.

The target was of course to achieve the maximum compression with minimum deterioration of intelligibility. The naturalness of the synthesized speech was a second order criterion for defining the speech quality.

The conclusion of this research is that a synthesized speech with more than 75 percent syllable intelligibility can be achieved using a transmission rate of 3750 bits/sec. The compression ratio is 15 to 1 and the naturalness is evaluated to be between satisfactory and good. The optimum block length was found to be 16 msec (128 samples) with selection of $1/32$ of the transform components and using 4 bits only for amplitude quantization.

The local maxima method utilizes digital transmission of the compressed signal, which is an additional advantage of this method. The parameters values can be changed so that standard transmission rates such as 9600, 4800, 2400 bits/sec can be used while the quality of the synthesized speech would change accordingly.

The crosscorrelation and mean square error between the compressed signal and the original one were calculated during the research, with the aim to find some relation between their values and the intelligibility. No match was found.

At the end of the thesis recommendations are given, for future continuation of this work, in order to improve the results.