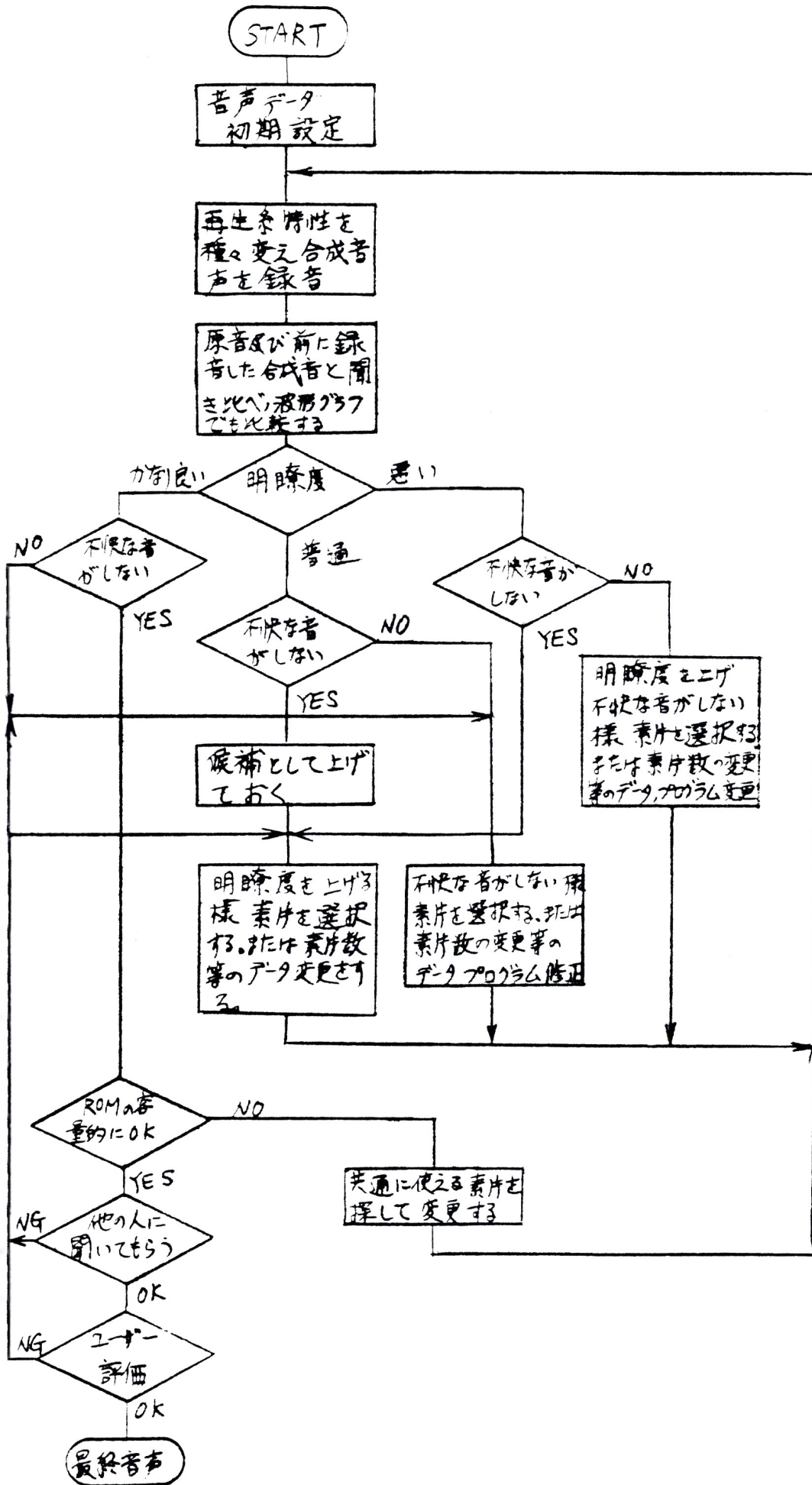


6.4 品質評価フローチャート



デモ用の音声品質 vs ROM 使用量 (例 こんにちは)

項目	音質		
	超高品質	高品質	中品質
トーン波形	1 (1132ワード)	1/4.4 (256ワード)	1/4.2 (80ワード)
エンベローフ	1 (48ワード)	1 (48ワード)	1/2 (24ワード)
基本周波数	1 (48ワード)	1 (48ワード)	1 (48ワード)
順序コントロール	1 (66ワード)	1/1.65 (40ワード)	1/3.7 (18ワード)
総合	1 (1294ワード)	1/3.3 (392ワード)	1/7.6 (170ワード)

本システムの音声品質 vs. ROM 使用量 (単位:ワード=16ビット)

項目	音質		高品質		中品質		低品質	
	男	女	男声	女声	男声	女声	男声	女声
100音節 トーン波形			4960 (16×3×100) +32×5	2560 (8×3×100) +32×5	3328 (16×2×100) +32×4	1728 (8×2×100) +32×4	1760 (16×1×100) +32×5	928 (8×1×100) +32×4
エンベローフ			256	256	256	256	256	256
基本周波数			256	256	256	256	256	256
1文字当りの パラメータ順序設定			5	5	3	3	2	2

音声合成 時間

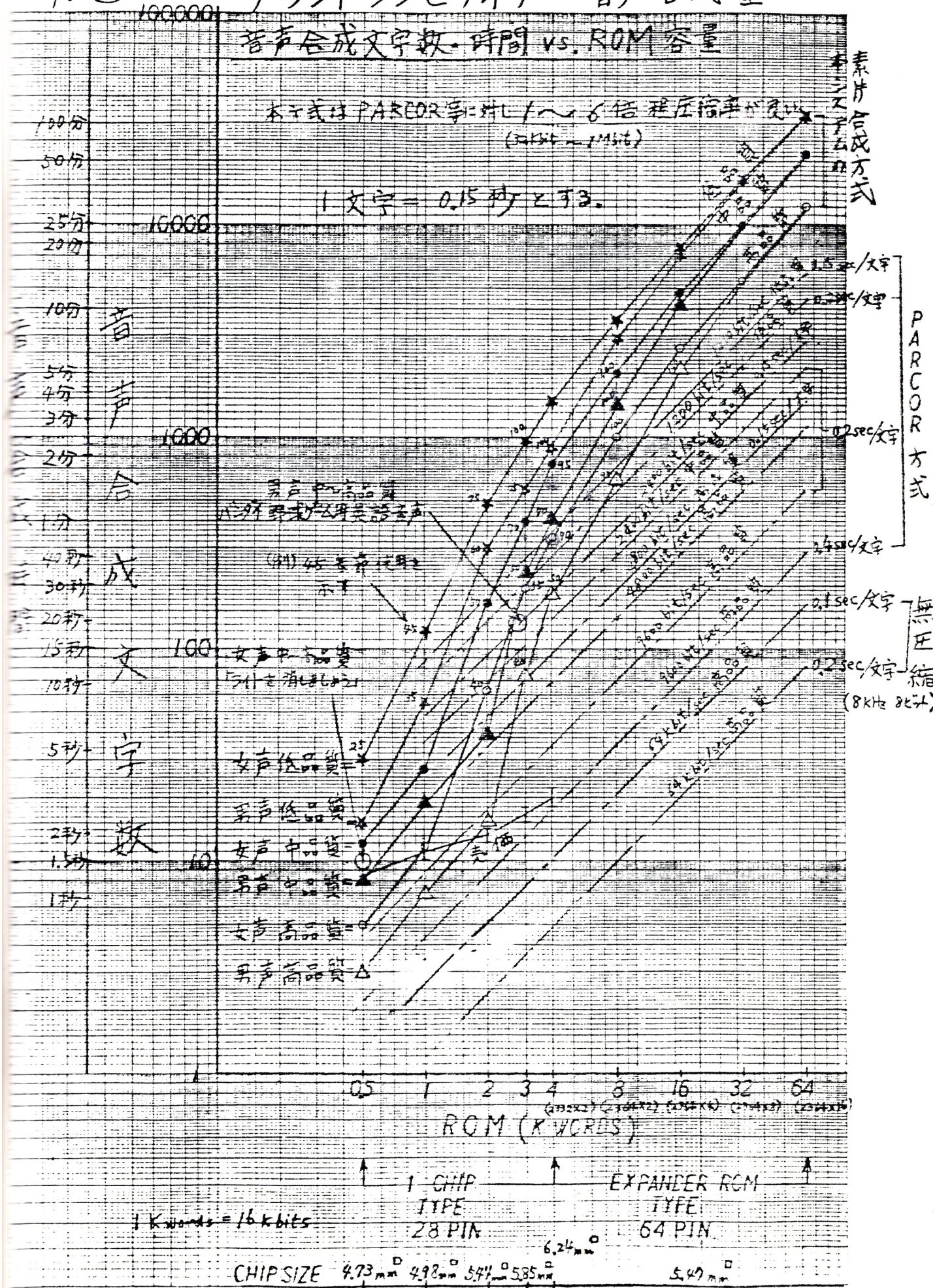
品質 品質	外付ROM最大 64 K word	1 chip 最大 3K word	
		女声	男声
低品質	80分	2分22秒	1分22秒
中品質	55分	1分	39秒
高品質	30分	29秒	11秒

1文字 = 0.15秒 とした場合

1.3

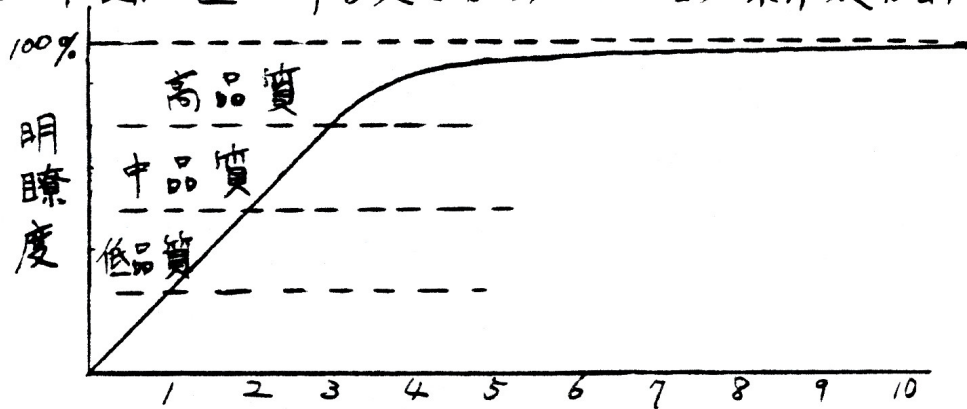
サウンドシンセサイザー音声合成量

音声合成文字数・時間 vs. ROM容量



6.5.5 音声素片数 と 品質

- 明瞭度を良くする方法としては 音節の特徴を良くあらわしている代表素片をデバッグしながら選択し、音声素片の数を1音節当り2~4波形位にする必要がある。音声素片数がある程度



1音節当りの音声素片の数

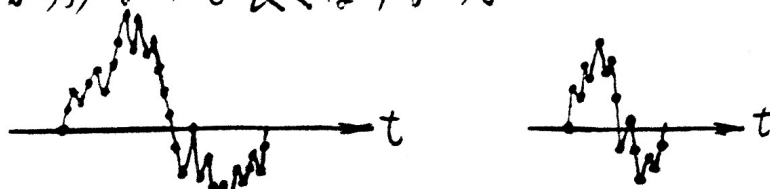
多くなると 明瞭度はそれに比例しては良くなりません。

「Wa」のような半母音は素片数を多くしないと品質は良くなりません。素片数が少なくても良い素片を選べば、明瞭度は良くなります。

6.5.6 ピッチ と 正規化音声素片分割数

- 男声より女声の方が基本周波数(ピッチ周波数)が2倍位高いので 同じ品質に対するデータの圧縮率が良い。

男声と女声のくり返し部の含有高調波成分(ホルマント)の上限の周波数帯域は同じなので 女声が1オクターブ高い分だけ1波形当りのサンプリング数は半分で済み 男声の1音声素片の分割数が64または32であれば女声のそれは32または16 となって、音声素片のROM使用量は1/2となり 女声のデータ圧縮率が男声よりも良くなります。



男声と低いピッチの女声...32分割, 女声(高いピッチの人)...16分割