

Internal Use Only (非公開)

TR-SLT-0078

テキスト音声合成ソフトウェア XIMERA 評価報告書
Evaluation Reports of Text-to-Speech Software XIMERA

津崎 実
Minoru TSUZAKI

河井 恒
Hisashi KAWAI

2004年8月20日

概要

本テクニカルレポートは、ATRの第3世代の音声合成ソフトウェア XIMERA(キメラ)に関する評価報告書あり、(1)「ベンチマーク報告書(音質編)」、(2)「資料：音質評価時の辞書登録単語」、(3)「ベンチマーク報告書(リソース編)」の3つの文書から構成されている。(1)は、XIMERAを含む11種類の日本語テキスト音声合成ソフトウェアを対象として行った音質評価実験の報告書であり、自然性評価実験およびアルファベット明瞭度試験の結果が解説されている。(2)は、音質評価実験の際にテキスト処理の誤りを回避するために辞書登録した単語の一覧である。(3)はXIMERAとCHATRを対象とした所要メモリ(RAM)、計算速度などの計算リソースに関する評価実験の報告書である。

音質評価実験の結果によると、XIMERAが評価対象中で最も高い自然性を有することが示されている。一方、リソース評価実験の結果によると、XIMERAの計算速度は、ハイエンドのPCを使用した場合、リアルタイム比がほぼ1であることが示されている。

(株) 国際電気通信基礎技術研究所
音声言語コミュニケーション研究所

〒619-0288 「けいはんな学研都市」 光台二丁目2番地2 TEL : 0774-95-1301

Advanced Telecommunications Research Institute International
Spoken Language Translation Research Laboratories
2-2-2 Hikaridai "Keihanna Science City" 619-0288, Japan
Telephone: +81-774-95-1301
Fax : +81-774-95-1308

©2004 (株) 国際電気通信基礎技術研究所

©2004 Advanced Telecommunications Research Institute International

はじめに

本テクニカルレポートは、ATR の第 3 世代の音声合成ソフトウェア XIMERA(キメラ)に関する評価報告書である。

ATR 設立以来現在までに 2 つの音声合成システムが開発されている。1 番目のシステムは、自動翻訳電話研究所(1986-1993)の研究成果として開発された ν -Talk であり、2 番目のシステムは、音声翻訳通信研究所(1993-2000)の研究成果として開発された CHATR である。XIMERA は、音声言語通信研究所(2000-2001)および音声言語コミュニケーション研究所(2001-2006)の研究成果として開発中のシステムであり、 ν -talk から数えて第 3 世代目にあたる。

XIMERA は波形素片接続方式に基づいており、その点では CHATR と同様であるが、ソフトウェアとしては全く関連がない。一方で XIMERA においてはテキスト処理、音声コーパス、韻律予測、素片選択評価関数などさまざまな点で改良が行われており、音質面では格段に進化している。

XIMERA は、一般的な入力テキストに対する自然性の向上を第一の目標として開発されているため、多言語、話者性、発話スタイル、感情など合成音の多様化への配慮はされていない。このため、対象言語は、日本語と中国語のみであり、発話スタイルは基本的に朗読調のみを想定している。

ソフトウェアとしての XIMERA は、主として「音声認識・音声合成マルチクライアントプロジェクト」の一環として、音声言語コミュニケーション研究所(SLT)と技術リエゾンセンタ(TLC)の共同作業により 2002 年度-2003 年度の 2 年間で開発された。それぞれの担当範囲は、SLT がアルゴリズムとプロトタイプソフトウェアの開発、および音質評価、TLC がソフトウェアのコード整理、品質管理、文書作成、および計算リソースの評価である。

XIMERA には、ソフトウェア開発の途上で作成された複数のバージョンが存在する。具体的には、ver-0.96, ver-0.97, ver-1.0, ver-1.1 の 4 つである。このうち ver-0.97 は開発用の内部向けバージョンであり、リリースはされていない(他のバージョンはすべてリリースされている)。「音声認識・音声合成マルチクライアントプロジェクト」における最終リリース版は、ver-1.1 である。開発版とリリース版の主な違いは、後者では中国語機能と日本語男声コーパスが取り除かれている点である。

本テクニカルレポートには、(1)「ベンチマーク報告書(音質編)」, (2)「資料：音質評価時の辞書登録単語」, (3)「ベンチマーク報告書(リソース編)」, の 3 つの文書が収録されている。(1)は、XIMERA を含む 11 種類の日本語テキスト音声合成ソフトウェアを対象として行った音質評価実験の報告書であり、自然性評価実験およびアルファベット明瞭度試験の結果が解説されている。音質評価実験においては、テキスト部より音声生成部の性能に重点をおいて評価するため、未登録語を辞書登録することによりテキスト処理の誤りを可能な限り回避している。その際に登録した単語の一覧が(2)の文書に記録されている。音質評価実験の対象とした XIMERA のバージョンは、聴取実験に約 1 ヶ月を要するという事情から最終版より少し古い ver-0.97 となっているが、ver-0.97 と ver-1.1 は素片選択・接続部のアルゴリズムが同等であるため、音質上の差はない。(3)は、所要メモリ(RAM)、計算速度などの計算リソースに関する評価実験の報告書である。評価対象のソフトウェアは、XIMERA(ver-1.1)と CHATR である。

音質評価実験の結果によると、XIMERA が評価対象とした 11 のシステム中で最も高い自然性を有することが示されている。一方、リソース評価実験の結果によると、XIMERA の計算速度は、ハイエンドの PC を使用した場合、リアルタイム比がほぼ 1 であることが示されている。

XIMERA (Ver1.0)

ベンチマーク報告書

(音質編)

2004年3月発行

2004年6月改訂

株式会社 国際電気通信基礎技術研究所
Advanced Telecommunications Research Institute International

目次

1.	目的と背景	3
2.	実験1	4
2.1.	方法	4
2.1.1.	刺激	4
2.1.2.	評定者	16
2.1.3.	手続き	16
2.2.	結果と考察	17
3.	実験2	28
3.1.	方法	28
3.1.1.	刺激	28
3.1.2.	手続き	28
3.1.3.	評定者	29
3.2.	結果と考察	29
4.	実験3	34
4.1.	方法	34
4.1.1.	刺激	34
4.1.2.	手続き	34
4.1.3.	評定者	35
4.2.	結果と考察	35
5.	全体考察	38

改訂の記録

2004年6月版では、主に次の修正を行いました。

- (1) “SMARTTALKER” → “SMARTTALK”
- (2) MOS 値の経時変化の記述を追加(pp.26-27)
- (3) 実験1と実験2の結果の相関関係についての記述を追加(p.33)

1. 目的と背景

XIMERAにより合成される音声の品質がどの程度のレベルに達しているかを調べるために、XIMERAならびに市販のTTSシステムとの比較評価実験を行った。評価実験の場合、合成音を作成する入力テキストに依存して品質が左右されることは経験上よく知られている。また、評価といってもそこには各種側面が存在することは直観的に分かる。例えば入力テキストとは異なる読みを与えている一方で読み上げられた音声の自然性は高いという場合と、入力テキスト通りに正確に読み上げているものの自然性が低いような場合で、どちらが良いかの判断はそれぞれの側面にどう重みを付けるかによって異なってくる。

日本電子工業振興会が制定している「音声合成システム性能評価のガイドライン」では、それぞれの要因についてSD法に則り別の評価を求めるやり方を評価の指針としているが、その一方で総合評価の方法については明確な指針を与えることが難しいとしている。

今回は、評価者に特に判断の根拠を明示することなく、漠然と読み上げた音声の良さを評価させるという方針で実験1を行い、それについての補充として、よりアプリケーション場面を想定した実験2、音節明瞭度により焦点を当てた実験3を実施した。実験1の場合、評価者によって判断の拠り所が異なってしまい、得られた結果によって優劣を議論できなくなるのではないかと、という懸念をされる向きもあるかもしれない。しかしながら、現代の統計手法はこのような評価者による違いを乱数要因と見なして、それを誤差項として統計的な検定をかけることを可能としているので、評定者によるバラツキを超えて有意な差があるかどうかを決定することができる。そのために、40名という十分な数の評定者と、10分野それぞれ10文という十分なバリエーションを持つ入力テキストを使用することとした。

2. 実験1

実験1では総合的な良し悪しについて絶対判断をする課題とした。各試行単独提示される合成音声について7段階の尺度によってその良し悪しを判断することを評定者に求めた。

2.1. 方法

2.1.1. 刺激

実験刺激は表1に示す11種類の合成器によって合成した。どの合成器も女性音声による合成とした。合成器によっては話速や平均F0などの各種パラメータを変更可能なものもあり、その調整により出力音声の音質が変化する可能性が存在した。今回の合成にあたってはそれらの調整値は基本的に各合成器を提供しているベンダーによって設定された既定値をそのまま採用することとした。従ってそれぞれの合成器にとって最良のものであることを独自に検証したことにはならない。しかしながら、各ベンダーがわざわざ品質の劣る設定を既定値としているとは常識的に考えにくく、既定値を採用することによって大凡それぞれの最良の音質を得ていると仮定できる。

表1 合成器の種類と供給元

番号	製品名	バージョン	開発/発売元	API	話者名
0	XIMERA		ATR	独自	
1	CHATR98		ATR	独自	
2	ボイス君のテキスト スピーチ	V1.0.1.0	NTT データ	独自	
3	smartvoice4 xp		NEC	SAPI5	NEC Female
4	SMARTTALK for WindowsV3	V3.01	沖電気	独自	
5	ViaVoice for win standard	R10.0J	IBM	SAPI4	日本語・標準 Hanako (女性)
6	LaLavoice2001		東芝	SAPI4	お姉さん
7	ドキュメントトーカ Ver3.5	V3.5	クリエートシス テム	SAPI4	花子
8	Visual Speech Creator		NTT アイティ	独自	
9	Fine Speech	V2.1	アニモ/富士通	SAPI5	ケイコ
10	RealSpeakV2		スキャンソフト	独自	

入力するテキストとしては文体の偏りをなるべく避けつつ、かつ比較的現実的な応用場面を考慮した10種類ジャンルについて、それぞれ10文を準備した。10種類のジャンルは

表2に示す。各ジャンルの文例は表3に示す。尚、表3には各ジャンルごとに20文が示されている。これは当初の計画では20文を用いる予定であったものの、実際にはテキストから合成音声の作成がすべて手動によらなければならない合成器が存在したために、刺激作成に要する時間と稼働を抑えるべく表3に示す文の偶数番号のものだけを用いた。

表2 入力テキストのジャンル

ジャンル記号	ジャンル名
A	旅行対話
B	ATR503文
C	ニュース
D	コールセンター
E	構内放送
F	カーナビ・交通情報

表3 ジャンルごとの文例

A	旅行対話
A-01	トラベラーズチェックは使えますか。
A-02	ここで写真を撮ってもいいですか。
A-03	タクシー乗り場はどこですか。
A-04	日本語のパンフレットはありますか。
A-05	チェックアウトをお願いします。
A-06	ルームサービスをお願いします。
A-07	クレジットカードは使えますか。
A-08	わかりました、サービス料は含まれていますか。
A-09	何か申告するものはありますか。
A-10	タクシーを呼んでください。
A-11	モーニングコールをお願いします。
A-12	パスポートを見せてください。
A-13	どのくらい時間がかかりますか。
A-14	おねがいします、もっとゆっくり話してください。
A-15	いいえ、もっと安い部屋はありませんか。
A-16	はい、メニューを見せてください。
A-17	ここに座ってもいいですか。
A-18	すみません、予約の再確認をしたいのですが。
A-19	救急車を呼んでください。

A-20	ありがとうございます、おつりはとっておいてください。
B	ATR503 文
B-01	後手にまわって負けた警察が奮起する番だ。
B-02	今後ほぼ全面的に行政が怠慢になる恐れがある。
B-03	ふりあおぐと、すぐ頭上を、光が走った。
B-04	一般論でいえば、援助する側は、援助される側の屈折した心理を見落としがちである。
B-05	私の指には宝石の指輪はもうはめられません。
B-06	自分の好きなジャズピアニストの演奏への、震えるような感動を綴る人もいた。
B-07	この意見に宇宙人たちは、パチパチと、それぞれすべての手で拍手して賛成した。
B-08	小さな家に引っ越し、はや九ヶ月が過ぎさりました。
B-09	不思議な感銘と不意を突かれる驚異を覚える時、感情移入せずにはいられなかった。
B-10	また、襟や袖口、ポケット口などが油汚れで変色をおこすこともあります。
B-11	はじめは喜んだ見舞客をいやがるようになった。
B-12	私は上着を脱ぎ、石組みの上に両手をついて、うつぶせになった。
B-13	私たち夫婦は、軍の依頼で中国人運転手の通訳をしながら上海まで逃避行の同伴をした。
B-14	円高ショックを少しでも和らげようと自動車、電気などの機械産業全般に広がった。
B-15	最後まで、どうも不思議なほどおいしいなと思い続けた。
B-16	これほど躍動感がなくて、静かで沈痛な胴上げというのも珍しい。
B-17	むろん、これは偶然の一致ということが大いにある。
B-18	医院を出たあと、気づいて腕時計を見るとちょうど五時半であった。
B-19	一週間ばかりニューヨーク取材した。
B-20	たばこの焼け焦げは、見苦しいものです。
C	ニュース
C-01	今回の大統領の訪日を局面打開の絶好の機会として捉えたいと述べた。
C-02	今後の株価の動向を見据えた上で、迅速かつ有効な対策を講じていく必要がある。
C-03	本日の未明、関東地方を中心に強い揺れを観測しました。
C-04	厚生労働省は14日、日本を含むBSE発生国産の牛の背骨が付いた食肉の販

	売や、背骨を原材料とした食品、添加物の製造と加工を禁止することを決めた。
C-05	道路公団の総裁人事を巡っては、公団が債務超過であることを示す財務諸表が隠ぺいされるとされる問題などに関し、伊藤前総裁の対応が不誠実だったとして、原田国交相が10月24日に伊藤氏を解任、後任選を進めていた。
C-06	東京都世田谷区の信号付き交差点で昨年2月、乗用車を運転中、バイクの男性をはねて重傷を負わせたとして業務上過失傷害罪に問われた都内の主婦に対し、東京地裁は13日、無罪を言い渡した。
C-07	決勝トーナメントは8日から始まる。
C-08	シンガー・ソングライターの小城美紀さんが9日に東京都内に住む会社員と結婚していたことが11日分かった。
C-09	14日午前の東京外国為替市場の円相場は、上半期の国内総生産実質成長率がプラスになったことなどを受けて円買いが優勢となり、1ドル、引き続き108円台前半の円高、終値となった。
C-10	補佐官は米国の対イラク政策について「日本から全面的な支援を得ていると受け止めている」と指摘した上で「何が出来るかは日本が自ら決めるべき事柄だ」とも述べ、日本政府の対応を見守る考えを強調した。
C-11	県警で家族らから事情を聞いたところ、進路や成績など学校生活について悩みを抱え、登校を巡って毎朝のように家族ともめていたという。
C-12	《姓》陣営は選挙戦終盤に劣勢が伝えられており、県警は苦戦で追い詰められたための犯行とみて追及している。
C-13	14日の東京株式市場は3日ぶりに大幅反落し、日経平均株価の終値は前日終値比170円61銭安の1万167円06銭と、8月19日以来の1万100円台で取引を終えた。
C-14	中盤まで大きくリードした第3セットを失うなど、要所要所でミスが出て最後まで波に乗れなかった。
C-15	ロイター通信によると、車のすぐ前を米軍車両が走っており、米軍を狙ったイスラム武装集団のテロ攻撃の可能性が強い。
C-16	扇形に見えるのは南半球のクレーターの中にあり、長さ13キロ、幅11キロほどで、数十億年前にできた地形とみられる。水が流れたような跡が複雑に入り組んでいる。
C-17	上野の都美術館で開かれている「大英博物館展」は、15日から土日と休日の閉館時間を1時間延長し、午後6時までとすることにした。
C-18	まん延防止策では、水域全体の汚染が懸念される霞ヶ浦などでのコイの移動禁止や出荷自粛を継続する。
C-19	米国で中枢同時テロが起き、北朝鮮による日本人拉致の実態が明らかになるに

	つれ、そうした主張の非現実性を多くの国民が認識するようになった。
C-20	一方、BBCテレビは日本時間10日午前、日本の総選挙のニュースを流し、自民党を中心とする連立与党が過半数を維持したものの、改選前議席数には至らなかった、と報じた。
D	コールセンター
D-01	お待たせいたしました。お客様のお名前をお願いします。
D-02	製品の登録番号を教えてくださいませんか？
D-03	申し訳ございません。しばらくそのままお待ち下さい。
D-04	最新バージョンは当社のホームページ、《URL アドレス》からダウンロードできます。
D-05	恐れ入りますが、ご住所と電話番号、電子メールアドレスをお教え下さい。
D-06	どのようなお問い合わせでしょうか？
D-07	かしこまりました。後ほどこちらから連絡させていただきます。
D-08	今回はご迷惑をおかけして、大変すみませんでした。
D-09	あいにくと担当者の《姓》がはずしておりますので、10分ほどしてからもう一度お電話頂けますでしょうか？
D-10	確認いたします。《市名》市にお住まいの《姓名》さまでいらっしゃいますね。
D-11	お客様の使用されている環境についてお伺いしてもよろしいでしょうか？
D-12	エラーメッセージの内容について教えてくださいませんか？
D-13	当社の製品のご使用は今回が初めてでしょうか？
D-14	次のどの製品についてのお問い合わせか、選んで頂けますか？
D-15	弊社とのサポート契約は結んでおられるでしょうか？
D-16	繰り返します。《月》月《日》日、《駅名》駅、《時刻》時《分》分発、《列車名》、禁煙席を大人《数量》人のご予約ですね。
D-17	ようこそ、《人名》様。いつも《社名》をご利用頂きましてありがとうございます。
D-18	発信音の後に、チケットの枚数をボタンで押して下さい。
D-19	お客様がおかけになった電話番号は現在使われておりません。新しい番号を申し上げます。新しい番号は《電話番号》です。
D-20	あいにくとお客様の登録が確認できません。赤いチェックの付いた項目について再度ご確認の上、もう一度ログインして下さい。
E	構内放送
E-01	お客様のお呼び出しを申し上げます。《姓名》様、《姓名》様、いらっしゃいま

	したらお近くの係員までお申し出下さい。
E-02	まもなく1番線にまいります列車は《時刻》時《分》分発の「列車名」新大阪行きです。皆様、白線の内側でお待ち下さい。
E-03	駆け込み乗車は大変危険ですので、お止め下さい。
E-04	電車とホームの間が空いている場所がございますので、足下には十分ご注意ください。
E-05	この電車は各駅停車《駅名》行きです。《駅名》には後の急行が先着いたします。
E-06	只今3番ホームに停車中の電車は、《駅名》行きの急行電車です。途中、《駅名》、《駅名》に止まります。お乗り間違えのないようご注意ください。
E-07	只今、駅構内改装工事に伴い、本日の《駅名》行き列車は4番線からの発車に変更となっております。
E-08	今白のシャツ、赤のチョッキに、グレイの半ズボンの、《氏名》くんという2歳の坊やを探しております。
E-09	ピンクのセーターに、白のスカートの、《氏名》ちゃんという3歳の女の子が迷子になっております。
E-10	7階、レストラン、美術画廊、工芸品のフロアでございます。
E-11	3階、婦人服、呉服、アダルト・レディース・ファッションのフロアでございます。
E-12	5階、紳士服、鞆、スポーツ用品のフロアでございます。
E-13	本日のお買得商品は、青森産リンゴ、北海道産男爵いも、南米チリ産ブラックタイガー海老です。
E-14	3時よりのタイムサービスとして、お肉コーナー、当店平日価格の2割引でご奉仕させていただきます。
E-15	本日は8時30分を持って閉店とさせていただきます。閉店時間までごゆっくりお買い物をお楽しみ下さい。
E-16	お手元の整理券をお確かめの上、アルファベットの整理券記号別に列にお並び下さい。
E-17	館内での喫煙は、1階入り口横の喫煙コーナーにてお願い申し上げます。
E-18	本日、7階、催し物コーナーにて《イベント名》を行っております。皆様のお越しをお待ちしております。
E-19	当店、歳末謝恩セールとして年内は休まず営業いたします。営業時間は平日午前10時から午後8時、土曜、日曜、祝日は午前9時から午後7時とさせていただきます。
E-20	只今、3階婦人服売り場にて赤のカーディガンをお買いもとめいただきましたお客様、ご連絡したいことがございますので、3階婦人服売り場までお越し下

	さい。
F	カーナビ・交通情報
F-01	次の交差点を左折です。
F-02	まもなく《右・左》方向です。《施設名》が目印です。
F-03	設定したコースからはずれました。このまま誘導を続けますか？
F-04	まもなく目的地付近です。
F-05	《ガソリンスタンド名》のガソリンスタンドが目印です。
F-06	約1キロ先、分岐点です。そのまま直進して下さい。
F-07	600メートル先、大きく右方向、続いて左方向です。
F-08	交通情報を受信いたしました。
F-09	次の角を右折してください。
F-10	ルート上、首都高3号線に渋滞が発生しましたので、一般道への迂回経路を選択します。次の用賀出口で一般道へ降りて下さい。
F-11	信号機手前、右手にファミリーレストランのある交差点を左折です。
F-12	あと《距離数》キロほどは道なりに走行して下さい。
F-13	明治通り、《地名》付近に故障車がいるため上下線とも混雑しております。
F-14	首都高速六号線、箱崎ジャンクションを先頭に上り方向5キロの渋滞です。
F-15	阪神高速5号湾岸線、住吉浜入口、リフレッシュ工事のため夜間封鎖中です。一般道へお回り下さい。
F-16	経路上、後20キロほどでガソリンスタンドがあります。経路をはずれても構わないようでしたら、ここから1キロのところに最寄りのガソリンスタンドがあります。
F-17	目的の《施設名》までの予想所用時間は《時間》時間《分》分です。途中渋滞しそうな箇所はありません。誘導を開始しますか？
F-18	目的地までの有料道路利用料金は合計で《金額》円となります。この経路でよろしいですか？
F-19	最短経路には道幅の狭い区間があります。広い道を通るルートを選んだ場合は5キロ程度長くなります。最短経路を選びますか？
F-20	燃料が残り少なくなっています。次のサービスエリアを過ぎると、その次のサービスエリアまでは49キロです。
G	天気予報
G-01	《県名》県《東西南北》部、今日は晴れ時々曇り、比較的暖かい一日となるでしょう。

G-02	《県名》県《東西南北》部の明日の天気は、雨、降水確率は80%です。気温も高めで蒸し暑い一日になりそうです。
G-03	《県名》県《東西南北》部、来週は曇りがちの日が多くなりそうです。
G-04	本日の降水確率は、軒並み80%を超えています。お出かけの際には傘をお忘れなく。
G-05	《県名》県《東西南北》部は、週末にかけて好天に恵まれ、過ごしやすい陽気になりそうです。
G-06	このところ大陸から冷たい空気が流れ込んできており、朝晩の冷え込みが厳しくなっております。
G-07	本日は関東地方も初雪となりそうです。いよいよ本格的な冬の到来といえそうです。
G-08	今年は例年より桜の開花も早く、列島各地で春の到来を告げるニュースが飛び込んできています。
G-09	今日は典型的な五月晴れの一日となります。行楽には絶好の日和ですが、夕方から冷え込みますので、お出かけの方は何か羽織るものを持って行かれるのが良いでしょう。
G-10	ジメジメした天気が続きますが、来週の終わり頃には梅雨も明けそうです。
G-11	明日は日差しも強く、日中の予想最高気温も《温度》度となっており、いよいよ本格的な夏到来です。
G-12	大阪ではこれで熱帯夜が《数》晩続いたこととなります。
G-13	この厳しい残暑、まだまだ2週間ほどは続く模様です。
G-14	大型で並の勢力を持つ台風《数》号は、毎時《数》キロメートルで北北東に進路を取っております。
G-15	この晴天、明日までは続きますが、その後西の方から湿った低気圧が近づいてきておりますので、洗濯物は今日の内に済ましておくことをお勧めします。
G-16	この雨も今日の午前中であがる見込みで、明日の運動会は大丈夫そうです。
G-17	各地ともこのところ朝晩の冷え込みが厳しくなっております。
G-18	《地方名》管区気象観測台では昨日木枯らし一号を観測いたしました。
G-19	出発地付近では晴れておりますが、途中三国トンネルを越える辺りでは雪が予想されますので、タイヤチェーンなどの準備をお忘れなく。
G-20	三重県鈴鹿地方、レース開始時刻午後2時には小雨模様、その後1時間程度で晴れると予想されます。
H	緊急時情報
H-01	これは緊急放送です

H-02	只今の地階1号棟北端の火災報知器が作動いたしました。
H-03	只今の警報に基づいて現場の状況を確認した結果、特に異常は発見できませんでした。
H-04	火事です。速やかに係員の誘導に従って退去して下さい。
H-05	防火扉が後5分で閉鎖します。至急非難して下さい。
H-06	ガス漏れの可能性があります。火気の扱いに注意して、換気の上、ガス器具の点検を行って下さい。
H-07	現在、火口付近ではガスの噴出量が多くなっております。これ以上の接近は大変危険ですのでお止め下さい。
H-08	只今の震度は5弱です。停電のおそれがありますので、懐中電灯、携帯ラジオなどの準備をしてください。
H-09	先ほど発生した地震により津波の発生の危険があります。約《数字》分後に最高潮位に到達する予想ですので、指示があるまで高台に待避して下さい。
H-10	昨夜未明からの雨により、《河川名》が警戒水位に達するおそれがあります。引き続き警戒態勢を取って下さい。
H-11	2階北側バルコニーに侵入者を感知しました。
H-12	セキュリティーシステムを警戒モードにいたしました。
H-13	セキュリティーシステムを解除いたしました。
H-14	只今、非常に発達した雷雲が《地域名》上空を通過中です。外出の際は十分ご注意下さい。
H-15	《番号》号室の《姓名》さんの容態が急変しました。
H-16	《日にち》、《時刻》、《場所》で発生した車両事故が原因で只今《路線名》は《駅名》、《駅名》間が不通となっております。現在復旧の見通しはついておりません。
H-17	台風による河川の増水で、この《距離数》キロ先の《橋脚名》橋が危険な状態です。5キロ先の交差点で左折後、警察官の誘導にしたがって下さい。
H-18	只今、当ビルに危険物が仕掛けられたという情報が入りました。おそらく悪質ないたずらと思われるますが、念のためにビル内にいらっしゃる方には外へ待避していただきます。
H-19	今から《分》分前の《時刻》に、《住所》、《施設名》にて火災が発生、現在も炎上しております。化学消防車の支援出動要請の可能性がありますので待機して下さい。
H-20	本日、《時刻》頃、《住所》、《施設名》で発生した事故が原因で多量の有毒ガスが発生しております。風下にあたる《地域名》周辺の皆さんは指示のあるまで外出を控えて下さい。

I	娯楽情報
I-01	本日より《映画題目》を上映しております。
I-02	各回の上映開始時間は10時0分, 12時50分, 15時40分, 18時30分となっております。
I-03	当館では, 今後今年度カンヌ映画祭受賞作品を中心に, 珠玉の名作をラインナップしております。
I-04	シングルカップリング曲からも, 《人名》本人が楽曲をセレクト. さらに初回盤にのみ, ファンへの感謝を歌った新録曲をボーナストラックとして収録。
I-05	本日, 活きのいい天然物のハマチが入りました。《姓》様のご来店を店主一同お待ちしております。
I-06	当店の特徴はそのスープです。豚骨を強火で沸騰させるため, 骨の髄からゼラチンが溶けだし, スープと脂がトロリと混ざり合い, 一度食べたら忘れられない味です。
I-07	今回のツアーではバンド結成時のオリジナルメンバーでもある《外国人名》も特別参加。オールド・ファンは必見です。
I-08	《歌手名》《アルバム名》が2週連続1位を達成。先週までやや弱かったカラオケでも上昇し, 4要素とも上位につけてきた。
I-09	漫才コンビ「サイヤング」が11月23日深夜, 大阪の「浪速一番館」で初の単独オールナイトイベント「サイヤングのオールナイトギャグ」を行います。
I-10	ひとりで男と女, そして静と動を踊り分ける大曲に, 菊五郎が一世一代で臨みます。その脇を固めるのは鴈治郎, 吉右衛門, 仁左衛門という豪華な顔ぶれ。皆様, 是非ご来場下さい。
I-11	工場見学やイベント, 特別コースのご案内, 周辺のおすすめガイドなど, 盛りだくさんの情報をお届けします。見学は無料。どうぞお気軽にお越し下さい。
I-12	この特別展では, ルーヴルのピラミッド建設以前に, ナポレオン広場で1984年から1986年にかけて実施された考古学発掘作業の結果に基づきルイ14世時代のパリの日常生活を紹介する。
I-13	恵まれた自然環境を生かしたりラックスできる空間の中で, "芸術とは何か"といった根源的なテーマのもと, 幅広い視野で企画される催し物などをおして, 芸術との出会いの場を提供しています。
I-14	当園の開園時間は午前9時30分から午後5時までです。入園券の発売と入園は午後4時までとさせていただきます。入園料は一般個人600円, 中学生200円, 65歳以上の方300円, 小学生以下は無料となっております。
I-15	ジャイアントパンダは中国のしせん省, かんせい省, に生息し, その数は現

	<p>在推定1000頭以下とされています。</p> <p>上野動物園では、1972年から飼育を始め、今までに3頭の繁殖に成功しています。</p>
H-16	<p>遠征軍で一度敵を倒した後に、再度遠征軍に参加した際モンスターを倒せないと、前回の活躍がリセットされてしまう仕様を変更しました。</p>
H-17	<p>プレイステーション用ソフトとしてスクウェアより発売された「ファイナルファンタジー・セブン」の続編「ファイナルファンタジー・セブンアドベントチルドレン」がDVD・ビデオで登場することが明らかになった。発売は2004年夏。ゲームではなく、映像作品となっている。</p>
H-18	<p>タイガーテイルは1999年4月3日生まれのフランス産です。生産者はスカイマーク牧場。調教先はフランスですが、これまでイタリア、英国、カナダでも出走し、遠征には慣れています。</p>
H-19	<p>金沢の冬の風物詩“雪吊り”に、伝統芸能の金箔をあしらったシンボルモニュメントが幻想的にライトアップされます。また、交差点近くの街路樹にも豆球のイルミネーションが施されます。</p>
H-20	<p>山の四季の彩りを見せてくれるこの橋は「くまなき月の渡るに似たり」という亀山上皇のお言葉からこの名になったとか。橋の下流の大きな楓は、11月下旬に茶色っぽくなり、散りそうになってから紅葉するので、お見逃しなく。</p>
J	<p>占い</p>
J-01	<p>一つの目標を持つことによって大きな幸運がもたらされる日でしょう。また、周囲の人たちと一致団結できそうな日です。そうした過程で、いつも近くにいる人の今まで気づかなかった長所を発見し、再評価することもあるでしょう。</p>
J-02	<p>交友関係を広げていくには良い時期なので、友人同士の集まりやパーティには積極的に参加してください。思いがけない場所で思いがけない人と知り合いになることがありそうなので、初めての場所でもどんどん足を踏み入れてみてください。</p>
J-03	<p>家族や親しい友人が、大きな幸運の情報や出会いをもたらしてくれることがありそうな日です。特に家族には、自分のやりたいことや欲しいものを日頃からアピールしておく、大きな援助を得られそうです。</p>
J-04	<p>ストレスが溜まって、周囲に当たり散らしてしまいそうな日です。そのせいでますます自己嫌悪に陥ったり、さらにストレスを抱えてしまうことにもなりそうです。まず自分の気持ちを冷静に見つめ直し、根本的な問題を見つけてから、ひとつずつ解決するよう努力してみてください。</p>
J-05	<p>現在、欲しいものがあるあなたは、親しい人から譲られたりたり、両親が力に</p>

	なってくれたり、いろいろなチャンスが訪れそうな日です。好きな人のいるあなたは、日頃から自分の気持ちを周囲の人にオープンにしておけば、交流のきっかけが得られるでしょう。
J-06	群れをなさない狼のように、人とは違う生き方を常に心がけているあなた。マイペースで何をするにも自己流がポイント。一人の時間や空間を大事にするので変わり者と思われがちだけど、変わってるといわれるのが嬉しかったりもします。
J-07	化けるのが上手なたぬきのように、どんなキャラにもなれるので、愛嬌があり上司などにかわいがられます。なぜか根拠のない自信を持っているから、何にでも積極的にチャレンジし、吸収するのが大得意。
J-08	俊敏な黒ひょうのようにスマートであることを大事にしているあなた。メンツやプライドにこだわり、リーダーシップを取りたがりです。しかもオシャレで新しもの好き。最新の情報・流行を敏感にキャッチして、人より1歩前を歩きます。
J-09	いつも群れをなして行動するヒツジのように、寂しがり屋で人懐っこく、いつでも周りに人がいます。和を大切にするので人脈作り・情報集めが得意。また、羊毛の下に本音を隠しているため、表面的には客観的かつクールに見えます。
J-10	異性のことはしばらく忘れて友情を深めよう。
J-11	平穏な一日になるでしょう。
J-12	ふとしたことでトキメキを感じそう。
J-13	たまったストレスを発散しましょう。
J-14	ビジネスチャンスがありそう。
J-15	周囲をよく観察してみましょう。
J-16	悩み事を相談してみましょう。
J-17	今日の買物に満足するでしょう。
J-18	今は自分に磨きをかけるときです。
J-19	大きな失敗をしてしまいそう。
J-20	空き巣には要注意。

テキストの作成は、「旅行会話」、「ATR503 文」を除いては実験者がインターネット上の関連ページの文例を土台としつつ、若干の修正を施して作成した。「旅行会話」については ATR が収集した旅行会話の対訳コーパス中から、出現頻度と文の長さのバランスを考慮して選んだ。基本的に出現頻度と文の長さには負の相関が見られるが、選定された文例はその中では出現頻度が高い割には文の長さが長いものである。ATR503 文はいわゆる音韻バラ

ンスのとれたセットである。今回はその中からのランダムサンプルした結果を用いた。

2.1.2. 評定者

評定者は健聴者 40 名で、全員が標準語圏で言語形成期を過ごし、かつ現在からさかのぼって 3 年以上標準語圏で生活者している者である。これはすべての合成器が標準語として自然と感じられるように韻律の付与を行っているためである。また、性別、年齢層の偏りを避けるために評定者の内訳は表 4 に示す通りとした。

表 4 評定者の性別と年齢層別の人数

年齢層	人数	男	女
20	13	7	6
30	14	7	7
40	13	6	7

2.1.3. 手続き

各評定者は各ジャンルにつきランダムに選んだ 8 文について、それぞれ 11 種類の合成器によって合成した音声刺激 880 個について、2 回ずつ判断を行った。従って 1 評定者当たりの総試行数は 1760 となる。各評定者は合成刺激の自然性に関して、

- 3 とても悪い
- 2 悪い
- 1 どちらかという悪い
- 0 どちらとも言えない
- +1 どちらかという良い
- +2 良い
- +3 とても良い

の 7 段階の評定尺度に従って判断した。この実験では受聴した音声は自然であっても読み間違いや読み付与の問題が生じている場合はマイナス要因として考慮するように評定者に教示した。この読み間違いの発生に関する判断をするための情報として、音声刺激の提示と同時にその時に読み上げている内容が文字列として提示された。なお、読み間違いについては、実験者が事前に一通りのチェックを行い、間違いを生じる場合のものについては単語登録を行うことによってその解消を図った。しかしながら、単語登録によっても間違いが解消されないもの、単語登録機能そのものが使用できないもの、単語登録によって別の箇所の問題が生じてしまったものについてはそれ以上の手当は施さなかった。

刺激音声の提示順は、引き続き 11 試行を 1 ブロックとした中には必ず 1 つの合成器による合成音声は 1 つずつ含まれる、という制約を除いてランダムとした。評定者には合成器が 11 種類あることについての情報は与えなかった。また同じ文内容が異なる合成器によって読み上げられた音声を連続して聞くことは偶然の発生の場合を除いてないことになる。従って、評定者にとっては各試行に出現する合成音声の読み上げる内容、ジャンル、合成

器の種類に関する予測は実質的に不可能な状況での実験であった。

実験の制御はラップトップパソコン (IBM ThinkPad G40) で行った。音声刺激はパソコンのオーディオ出力からヘッドフォン(SONY MDR-Z900)を経て評定者の両耳に提示された。平均的提示音圧レベルはA特性の重み付けで70 dBとした。また、通常のオフィスにいる環境に近い受聴環境とするため背景雑音をA特性重み付けの音圧レベルで48 dBにて実験中を通して提示した。背景雑音もヘッドフォンからの提示である。背景雑音は電子協騒音データベース No. 14 の計算機室のサンプルよりワークステーション1台による騒音を使用した。

実験は8つのセッションに分割して実施した。各セッションの本試行数は220試行となった。これに加えて11の練習試行を各セッションの冒頭に実施した。実験の進行のペースは各評定者に基本的には任せた。半数が終了した時点で2分間の強制的な休憩を挿入したほか、評定者の必要に応じて休憩を随時取ることを許した。従って1セッション当たりの所要時間は評定者ごとに異なってくるが、おおよそ1時間から1時間30分程度であった。各評定者は1日4セッションを上限として、複数日に渡って実験に参加した。

2.2. 結果と考察

図1は合成器のそれぞれに対する評価値の平均をプロットしたものである。もっとも平均点の高かったXIMERAに対して、最も低いドキュメントトーカーではおよそ2段階ほどの値の開きがあることが分かる。図2には入力テキストのジャンル別に各合成器の平均評価値を示したものである。合成器とジャンルを主要因、評定者をブロック要因とした分散分析を実施した結果、合成器の要因、ジャンルの要因、両者の交互作用とも有意差が観察された[それぞれ、 $F(10, 390) = 109.87, p < .0001$; $F(9, 351) = 100.56, p < .0001$; $F(90, 3510) = 58.61, p < .00001$]。Tukey-KramerのHSDによる合成器間の多重比較検定を実施した結果、最高評価値をとったXIMERAと次点のLaLaVoiceの間には有意な差が存在することが示された。この多重比較検定の結果は図1の中に線分によって示してある。棒グラフの上に線分で囲われた合成器の間には統計的な有意差が存在しないことを示す。

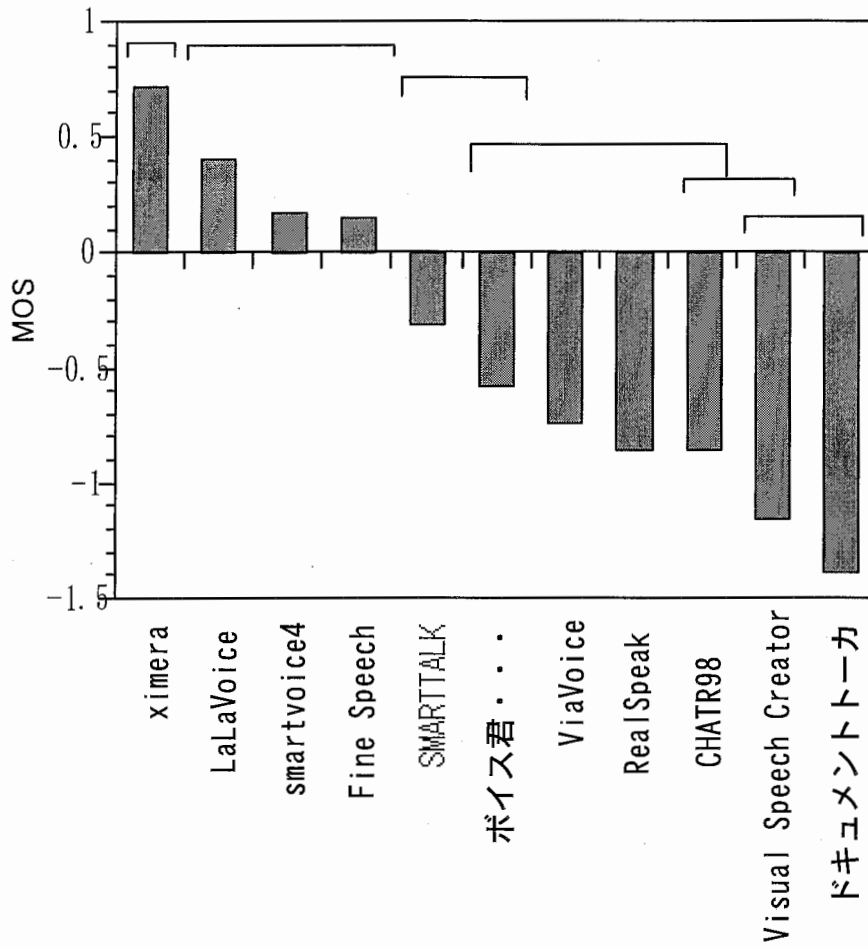


図 1 実験 1 に於ける合成器別の平均評定値. 各棒の上がひとつの線分で結ばれている合成器の間の差は統計的には有意ではない (有意水準 5%).

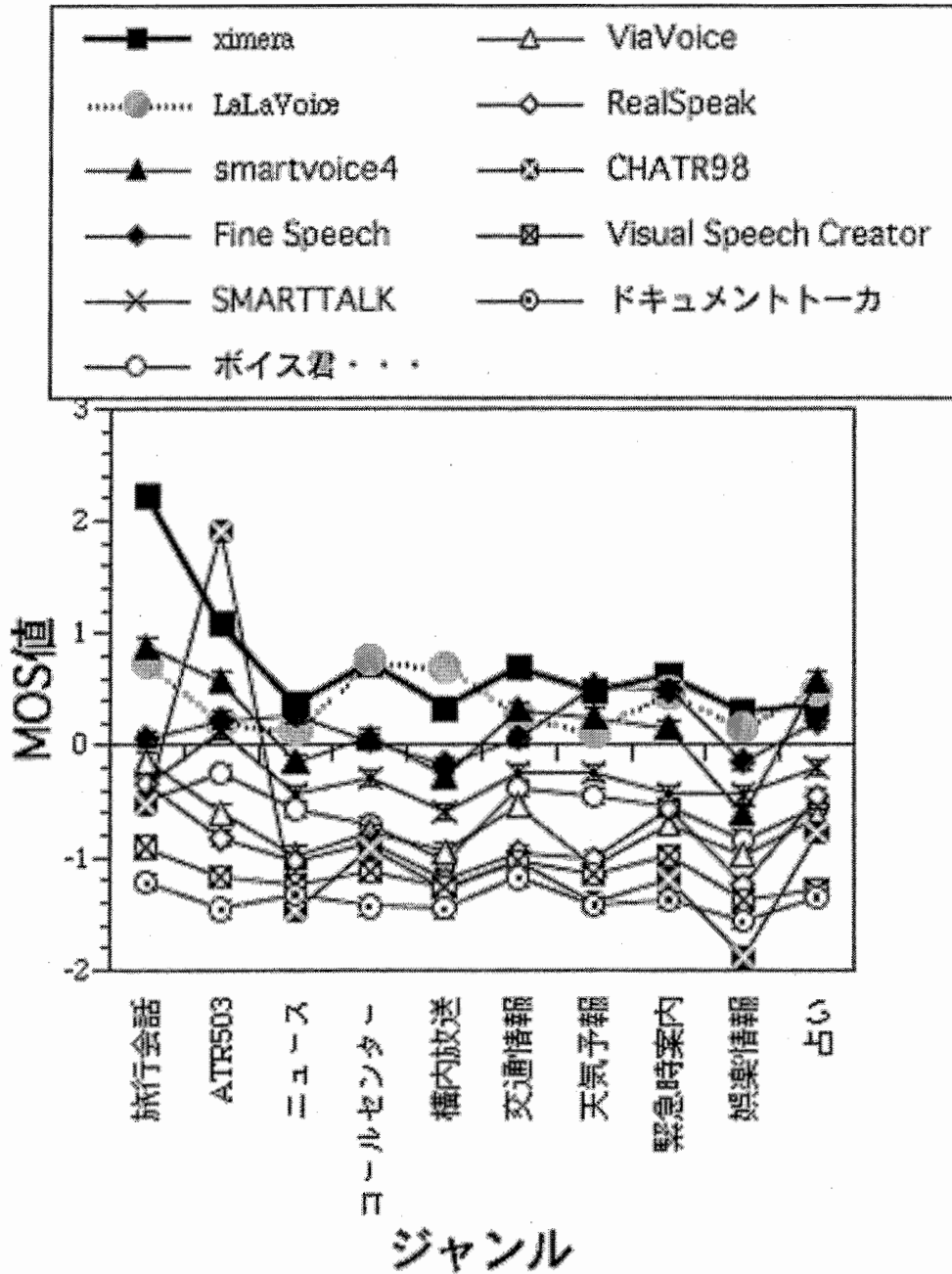


図 2 実験 1 に於ける評定の入力テキスト・ジャンルへの依存性。折れ線グラフの各線分は合成器の違いに対応する。

このように XIMERA は全体の平均を見る限り 2 番手の LaLaVoice に有意な差を付けたわけであるが、合成器とジャンル間に有意な交互作用が存在したことは、ジャンルによっては合成器の得手不得手が存在することを意味する。実際に図 2 を見てもジャンルが異なると、最も品質が高いとされる合成器が変化している様子がうかがわれる。もっとも顕著なのは ATR503 文においてそれ以外ではあまり評価の伸びない CHATR98 が最高値を示し

ているような例である。これは CHATR98 では ATR503 文を読み上げたコーパスを使用しており、このジャンルに関してはクローズド・テストに近い形となっているためと解釈できる。XIMERA についても最も優位性を顕著に示すのは旅行会話のジャンルであり、このテキストを読み上げた音声はコーパス中に存在している。実際に合成された音声のセグメント情報を見る限りにおいては、なんらかの形で接続は行っていることが確認されたが、同様の仕組みで評価が高くなっていた可能性は否定できない。図 1 に示すように全ジャンルを平均した評価で XIMERA が最高点を獲得しているわけであるが、その原因の大半が旅行会話ジャンルにおける高得点にある可能性が考えられる。これについて検討するために、旅行会話の条件を除いた残り 9 つのジャンルについて平均評点を算出した。その結果を図 3 に示す。旅行会話ジャンルを除いても XIMERA が最高評価を得ている。このデータセットに対して合成器を主要因、評定者をブロック要因とした分散分析を実施した。合成器要因は $F(10, 390) = 102.42, p < .0001$ と有意となった。下位検定として Tukey-Kramer の HSD による多重比較を行った結果では、XIMERA と LaLaVoice の間には有意差が存在しなかった。但し、3 位グループを形成する Fine Speech、smart voice 4 と XIMERA の間には有意差が存在し、これに対して LaLaVoice と 3 位グループとの間には有意差が存在していない。従って、旅行会話ジャンルを除くことによって XIMERA の優位性は若干の低下を示すものの、基本的には保たれていると考えられる。

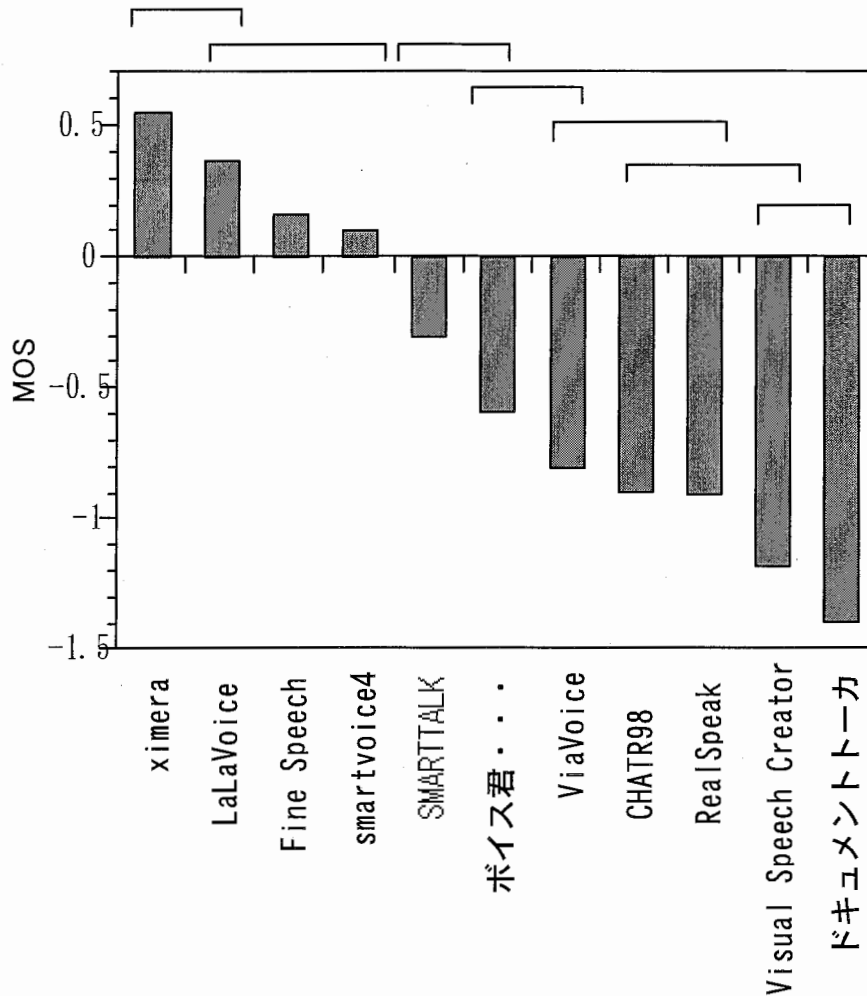


図 3 実験 1 において「旅行会話」ジャンルを対象外とした場合の合成器別の平均評定値. 各棒の上がひとつの線分で結ばれている合成器の間の差は統計的には有意ではない (有意水準 5%).

次に年齢層による評価の違いを見てみる. 図 4 は評定値を年齢群の関数として合成器別にプロットしたものである. 全体を通して右下がりとなる傾向がうかがわれる. これは即ち年齢の上昇とともに評価が下がっていくことを示す. 年齢層が高くなるにつれ合成音に対する違和感をより顕著に感じる傾向があるようである. XIMERA の持つ優位性は 30 代、40 代の群で際立っており、20 代では LaLaVoice や smart voice 4 と大差のない結果を示している.

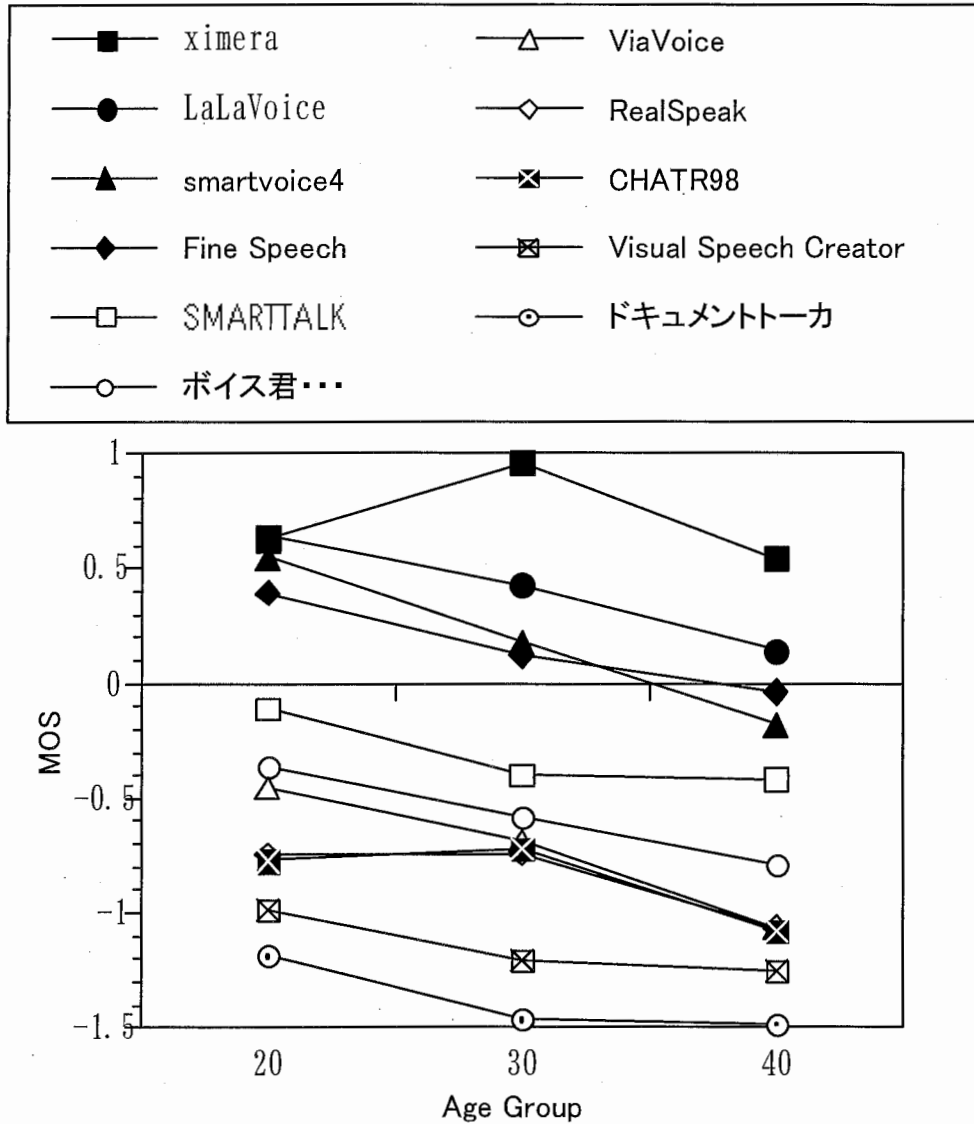
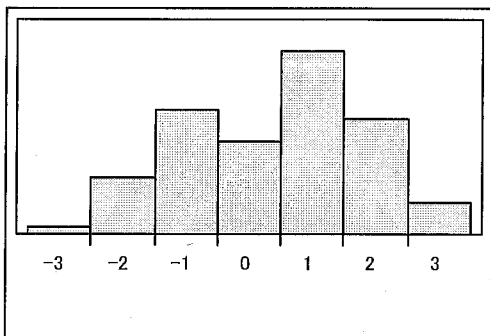


図 4 実験 1 に於ける評定の年齢層 (20 代, 30 代, 40 代) による違い. 折れ線グラフの各線分は合成器の違いに対応する.

実験 1 の結果についてさらに詳細に見るために、XIMERA と 2 位につけた LaLaVoice の間で与えられた評定値の分布を比較した。その結果を示したのが図 5 である。まず分かるのが XIMERA に対する評点では最高点(「非常に良い」)の 3 が与えられる頻度が LaLaVoice に比べて約 3 倍のオーダーで多くなっていることである。「良い」に対応する 5 点の頻度も LaLaVoice へ対する場合よりも高い。実験後に評定者に対して行った質問紙調査の回答の中には「自然音声と聞き分けがつかない程度品質の高いものが存在した」というものがあった。そのような印象を持ったときに最高点 3 が振られると考えると、XIMERA では全体の 17%ほどがそのカテゴリーであるのに対して、LaLaVoice は 5%程度に過ぎない。このように「非常に良い」「良い」の評価カテゴリーの選択率が上昇しているにもかかわらず、「非

常に悪い」「悪い」の評価カテゴリーの選択率は減少しておらず、むしろ若干上昇すらしている。要するに XIMERA の場合、うまく合成できるものは自然音声と比べても遜色のない程度に合成可能である一方で、問題を生じる場合は依然として許容しがたい問題点を残したままであると考えることができる。

synthesizer=lalavo

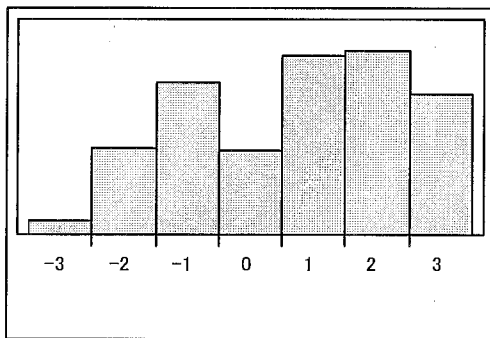


Frequencies

Level	Count	Prob
-3	92	0.01438
-2	587	0.09172
-1	1297	0.20266
0	971	0.15172
1	1911	0.29859
2	1202	0.18781
3	340	0.05312
Total	6400	1.00000

7 Levels

synthesizer=ximera



Frequencies

Level	Count	Prob
-3	103	0.01609
-2	665	0.10391
-1	1156	0.18063
0	636	0.09938
1	1365	0.21328
2	1404	0.21937
3	1071	0.16734
Total	6400	1.00000

7 Levels

図 5 実験 1 に於ける評定値の分布. (a) XIMERA に対するもの; (b) LaLaVoice に対するもの.

図6はXIMERAによる合成音の評価を入力テキスト別に平均したものである。誤差線は標準誤差を表す。平均点が有意に0よりも下となる例は100例中の14例に留まっている。表6にその文番号と平均値を示す。反対に有意に0を上回ったのは100例中68例と過半数を上回っている。

表6 実験1でXIMERAによる合成で有意に平均点が0未満となった文番号とその平均点

文番号	平均点
C17-0	-1.00
E13-0	-0.99
B01-0	-0.92
I03-0	-0.83
D09-0	-0.81
C11-0	-0.62
G01-0	-0.58
G19-0	-0.57
H17-0	-0.42
J07-0	-0.35
F17-0	-0.32
C19-0	-0.27
D17-0	-0.21
H13-0	-0.21

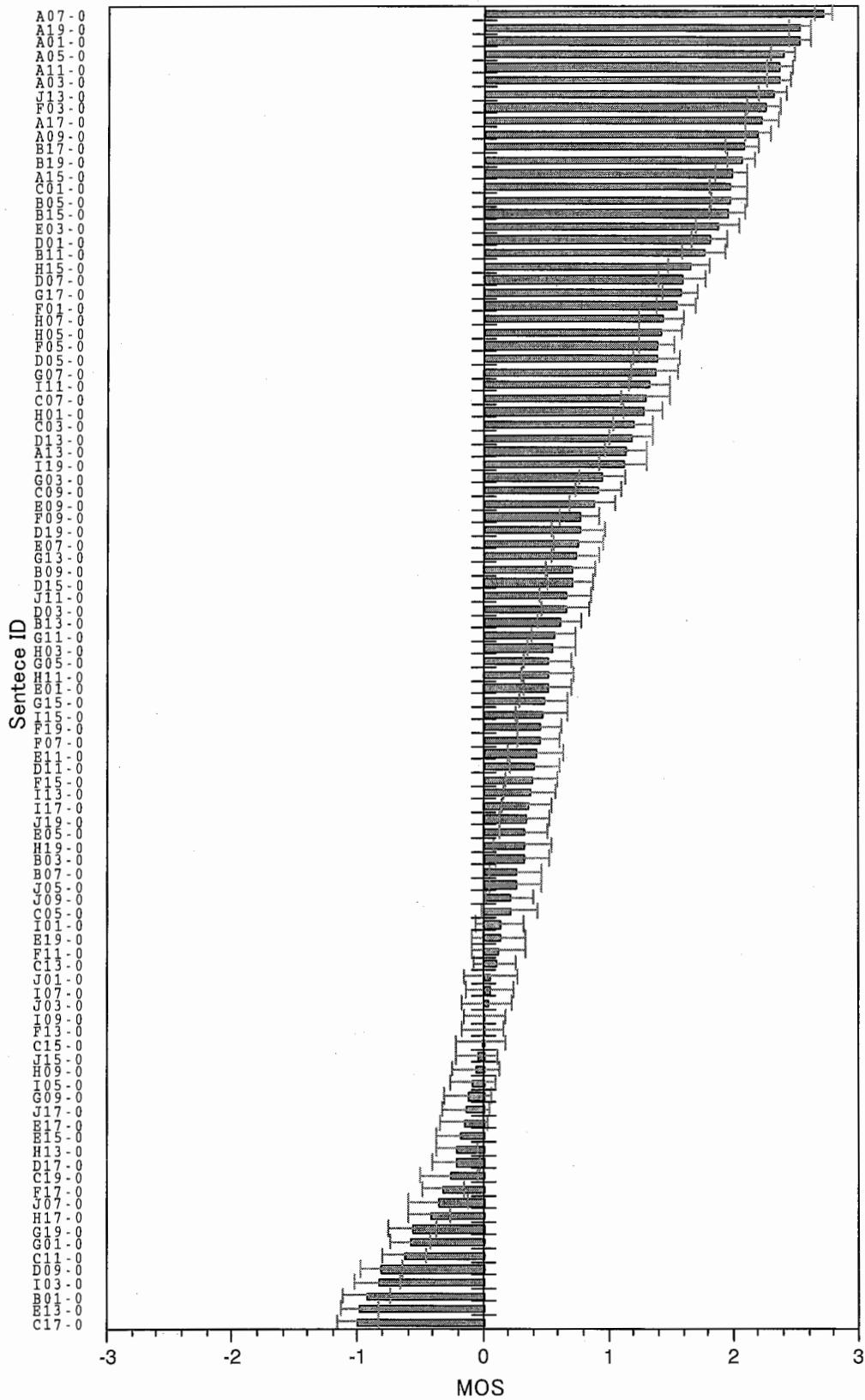


図 6 XIMERA の合成結果に対する入力テキスト毎の平均評定値 (誤差線は標準誤差)

一般論的な予測として、最初に聞いたときに感じた印象がくり返して聞くにつれ鮮明さを失っていくのではないかというものがある。このような現象が今回の評定実験についても生じていないかを確かめるために、各セッションにおけるそれぞれの合成器に対する評定の平均値を算出してみた。セッションの間には合成対象となる文の違いが存在しており、評定者毎にこの値を比較することは必ずしも合成音への暴露量に伴う変化と断定できないものの、施行する文の順番は評定者毎に無作為にその順番を変えており、評定者数も40名であることから被験者をプールした平均値によって暴露量による影響を見ることは妥当と判断できる。

図7はセッションの進行(即ち暴露量)の関数として見たMOS値の変化を、合成器毎に示したものである。この図を見る限り、例えば当初は顕著であった合成器間の差が暴露が増えるにつれて減少していくというようなことはない。セッションによる違いは実際には存在しているものの、特に規則的と考えられる傾向はない。敢えてそれを探すと、まず第1セッションと第2セッションでは総じて第2セッションの方が評価が落ち気味になっている。これはおそらく合成音に対する評価基準がこの間で若干変化したものであろう。第1セッションでは評定者の合成器に対する要求水準も定まっていない状態であるのに対して、第2セッションになると多少判定の基準を辛くしていくという傾向は一般論としては納得のいく傾向である。それ以降は、若干ではあるが評定値が上がっていく傾向が存在する。これはいわゆる慣れの効果と考えられる。ただ、このような慣れが起きた場合も評定者は合成器の間の性能差についてはなるべく維持して評価を下しているように今回の結果は見受けられる。

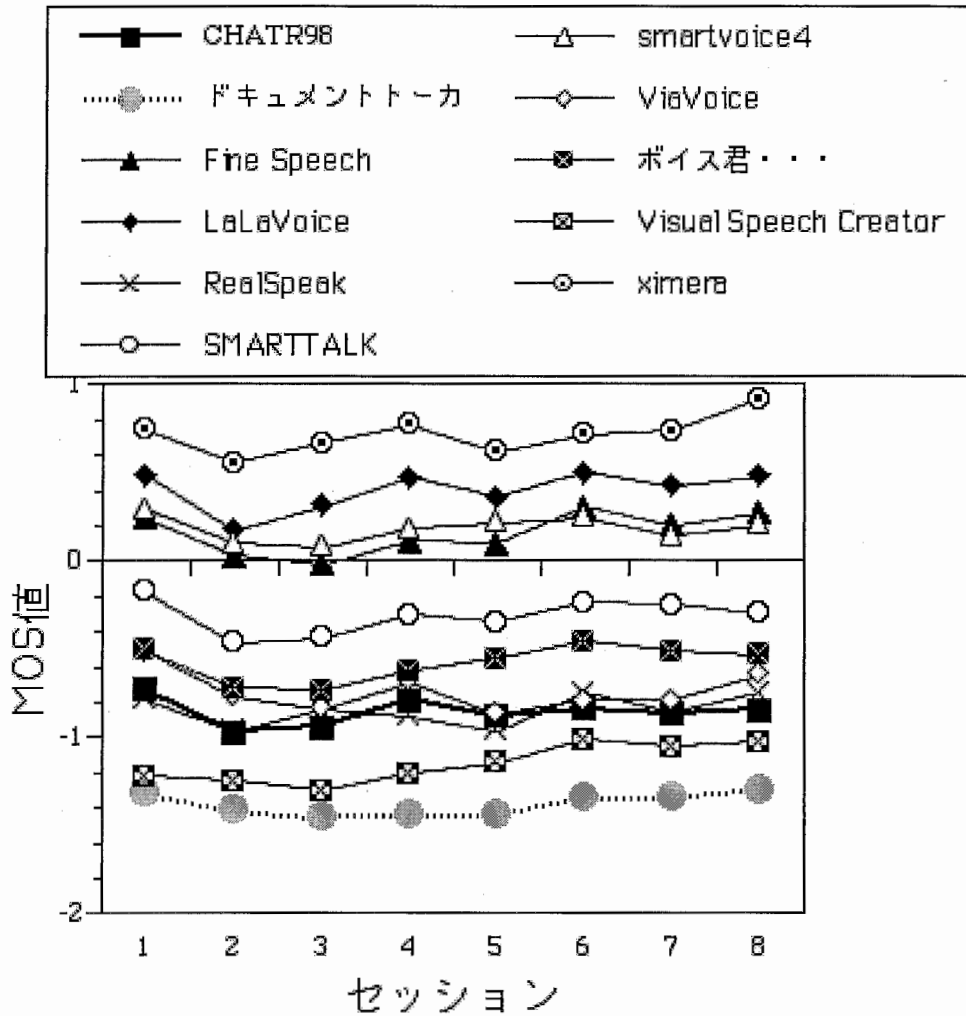


図 7 セッションの進行による MOS 値の変化

3. 実験2

実験1では読み間違いの存在を評定者に気がつかせる目的で読み上げているテキストを文字列として視覚的提示した。この方法の問題点は評定者がテキストに対する文字列情報を知ってしまうことにより、聞いたつもりになってしまうということである。トップダウン情報が人間の知覚を左右することはよく知られた事実である。実際に合成音を使う状況を考えるとテキストを視覚的に提示できない状況であるからこそ音声提示による恩恵が多くなるはずである。そのような状況では聞き手は自分がおかれた状況についての情報は持っているものの、どのような指示、案内が流れるかについては事前の情報を持っていないはずである。実験2では評定者に自分がおかれた状況なり合成音による情報提供サービスを受ける状況なりに関する説明を、各合成音を耳にする前に文字列として与えることにより現実的な状況を模擬して実験を実施した。

3.1. 方法

3.1.1. 刺激

実験1で用いた10ジャンルのうち、ニュース、コールセンター、構内放送、カーナビ・交通情報、天気予報の5ジャンルからそれぞれ2文を選択し、各合成器により合成した

表5 ジャンル別の選択文と状況説明

ジャンル	使用文	状況設定文
C	C05, C09	電車の中で携帯電話を使ってその日のニュースを聞くことができると想像して下さい。
D	D05, D11	電話による自動応答システムに問い合わせをしていると想像して下さい。
E	E05, E07	少し未来の駅やお店にいます。各所で携帯電話のあるボタンを押すと、受話器から様々な情報を聞くことができると想像して下さい。
F	F13, F15	将来型のカーナビゲーション・システムからの音声ガイドです。
G	G09, G11	新しいタイプのお天気情報サービスです。あなたの行動予定を教えておくと、その予定に関連した気象情報が受けられると想像して下さい。

3.1.2. 手続き

毎試行、状況説明文が視覚的に提示され、その状態でクリックすることによって合成音が提示されることとした。評定者は-2から+2までの5段階の尺度でその音声で情報サービスを受けることの善し悪しについての判断を求められた。その際にはサービスそのものの価値ではなく、あくまでも合成音声の品質に着眼した判断をするように予め教示した。5段階の判断基準として以下のような教示を与えた。

-2	【失格】	奇異さが気になり気が散るのでない方がいい
-1	【ギリギリ合格】	聞き流す程度でいいなら許せる
0	【合格】	実用的に許容範囲
+1	【優秀】	支障を感じない
+2	【理想的】	肉声と遜色ない

実験1同様、一連の11試行をブロックした中で各合成器が必ず1回ずつ提示されるという制約以外にはランダムとして提示された。各評定者につき、ジャンルと合成器の組み合わせ55組のそれぞれが4回ずつ提示された。評定者当たりの総試行数は220であった。これに11の練習試行を加えた計231試行を1セッション施行した。

それ以外の実験の詳細については実験1と同様であった。

3.1.3. 評定者

実験1に参加した40名の評定者が実験2に参加した。全員実験1を経験してから実験2に臨んだ。

3.2. 結果と考察

評定者、ジャンルをプールした合成器別の評定値の平均を図8に示す。XIMERAのみが「合格」点である0を上回る平均値を示していることが分かる。5つのジャンル別に各合成器の平均評定値を求めた結果を図9に示す。どこかひとつのジャンルでもいいから0点を上回る平均評点を獲得した合成器はXIMERA, LaLaVoice, Fine Speechのみとなった。このうち、XIMERAのみがすべてのジャンルで0点を超えている。実験1の時と同様に、合成器とジャンルを主要因、評定者をブロック要因とした分散分析を実施した。その結果、合成器の要因が $F(10, 390) = 62.79, p < .0001$ で、ジャンル要因が $F(4, 156) = 9.438, p < .0001$ で有意となった。両主効果の間の交互作用も有意であった[$F(40, 1563) = 11.84, p < .0001$]。下位検定として Tukey-Kramer の HSD による多重比較検定を実施した。図8に各棒の上の線分としてその結果を示す。最高平均評点を得たXIMERAとそれ以外の合成器の間の差が統計的にも有意となった。

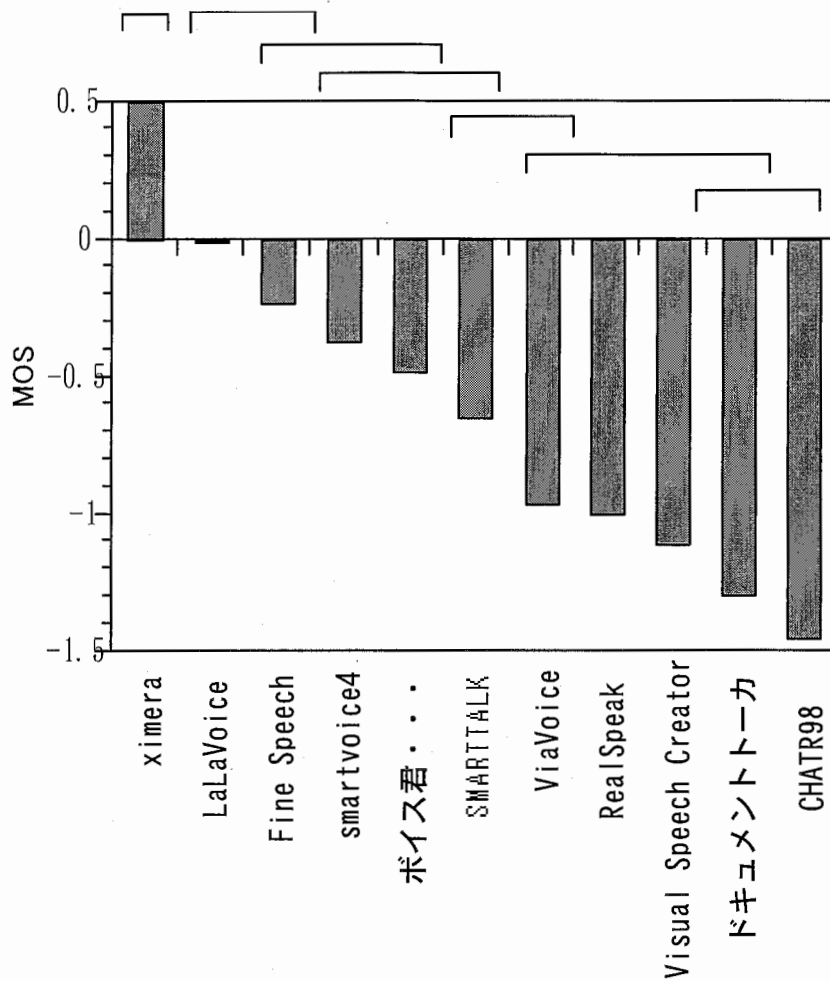


図 8 実験 2 に於ける合成器別の平均評定値. 各棒の上がひとつの線分で結ばれている合成器の間の差は統計的には有意ではない (有意水準 5%).

図 10 に年齢層別に平均評定値を示す. 実験 1 においては年齢が上昇するにつれて評価が低くなるという一般的な傾向がうかがわれたが, 実験 2 の場合はその傾向はあまりはっきりしない. いずれの年齢層においても平均点が 0 を上回ったのは XIMERA だけである.

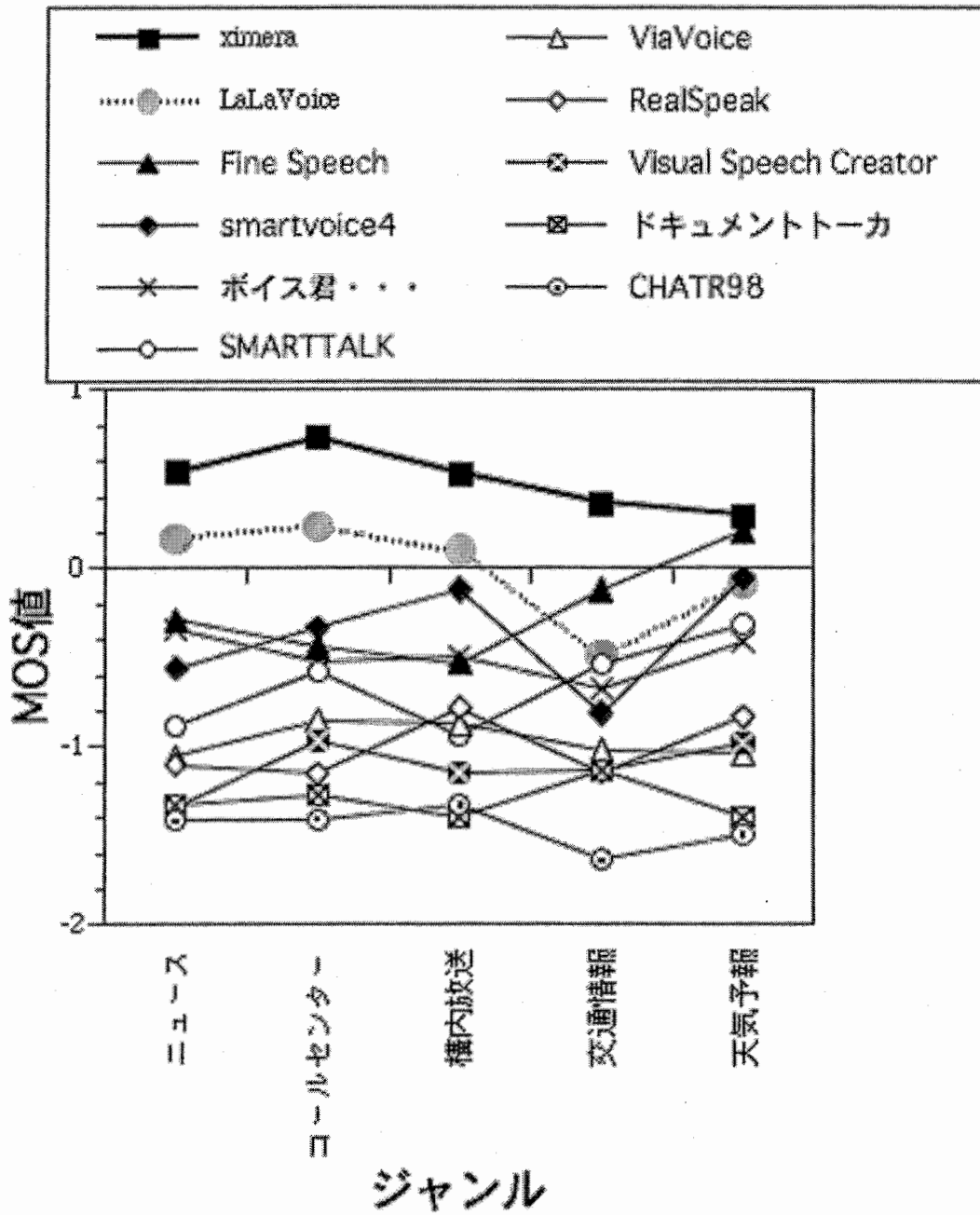


図 9 実験2に於ける評定の入力テキスト・ジャンルへの依存性. 折れ線グラフの各線分は合成器の違いに対応する.

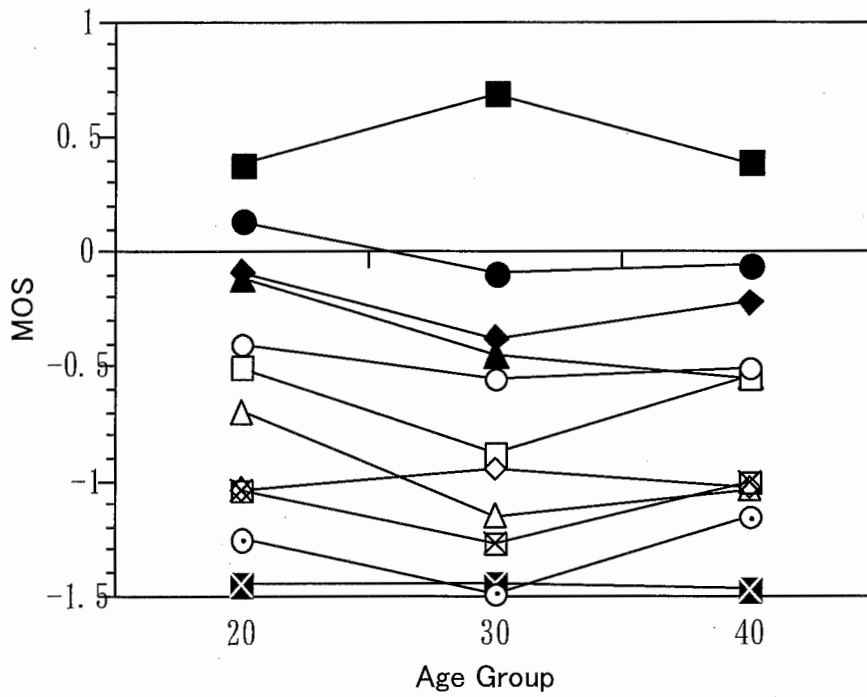
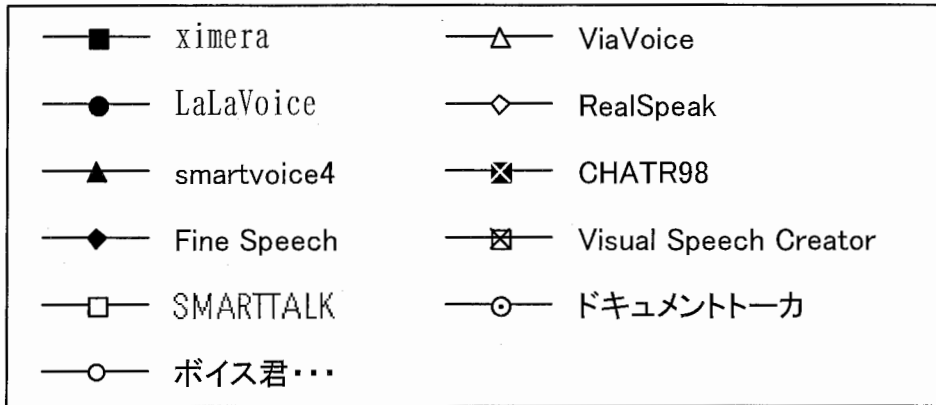


図 10 実験 2 に於ける評定の年齢層 (20 代, 30 代, 40 代) による違い. 折れ線グラフの各線分は合成器の違いに対応する.

実験1が特に使用環境を意識させずに漠然と合成音声の品質を問うたのに対し、実験2では使用される状況を思い浮かばせながらの実用性を意識した評価をさせた。この2つの方法の間に本質的な差が存在するか否かを検討する目的で各評定者、各合成器についての評定値の平均を2つの実験結果それぞれについて求め、両スコア間の相関を調べた。結果は図11に示す。相関係数は0.7744と高いものであった。この高い相関はふたつの実験の間に本質的な差が存在しないことを示す。

但し、はずれ値に関する分析をした結果興味を引く側面も見られた。図中赤線で囲んだ範囲は95%の信頼区間である。赤白抜きの丸印はximeraに対する評価である。青丸はximeraの次に良好な評価を得たLaLaVoiceに対する評価である。ximeraに対する点は2つの実験の間の回帰傾向にはずれた位置を示すものが多いことが分かる。そのはずれ方として実験1での評価に対して実験2での評価がことごとく高くなるためであることが見て取れる。

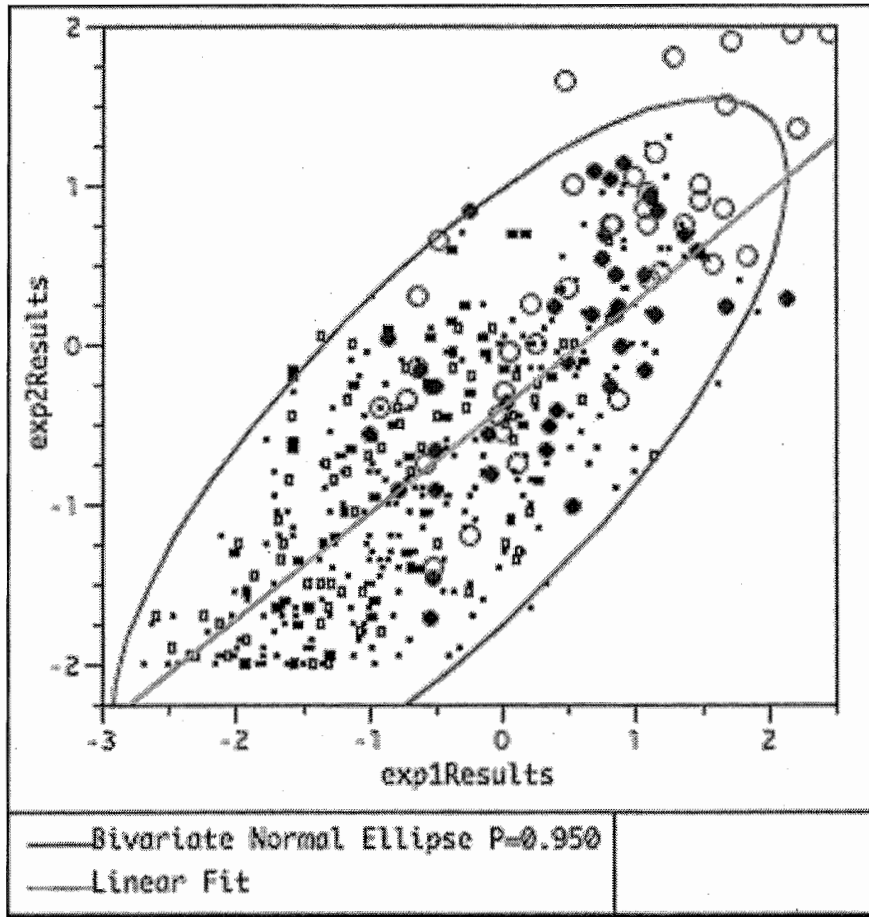


図 11 実験1と実験2の間の相関

4. 実験3

実験1と実験2では課題設定に多少の違いがあるが合成音を聴いたときの印象について調べた。このような印象評定の実験は、漠然としたイメージで判断することが可能であるため、得られた結果が本当に合成音声の品質に関わる指標を提供しているかどうか疑問視されることがある。今回の実験1のように評定者の人数、入力テキストともに多く、そのバランスも考慮されているような場合、上記のような懸念の多くの部分は回避できていると考えられる。しかし、より特定した合成音声によるサービスを前提とする場合、実験1, 2で検討を加えた以外の側面が重要となってくることも事実である。

通常、人間同士のコミュニケーション場面を想定しても音声会話は非常に冗長さの高い中で行われている。実験1や2での提示状況はそのような環境を前提として設計した。しかし、場合によっては冗長さの少ない形で情報伝達を音声を使用して行わざるを得ない場合がある。電話番号や電子メールアドレス、URLアドレスなどを伝える場合がそれに相当する。このような場合の情報伝達効率を調べることを目的として実験3は計画された。具体的には無意味なアルファベット列の読み上げを合成器を使って行い、それを聞き取るという課題による評価実験である。

4.1. 方法

4.1.1. 刺激

入力テキストはランダムなアルファベット列とした。ひとつのアルファベット列の長さは5文字から9文字とした。同一系列内に同じ文字が繰り返し入ることはなかった。アルファベットは26文字の中から無作為に選択されたが、合成器の制約により正しく読み上げられないものについては除外することとした。具体的には以下の3つの合成器についてそれぞれに示すアルファベットが除外された。

CHATR 98・・・"a" (「ア」と読み上げるため)

ボイス君のテキストスピーチ・・・"aiueon" (すべてローマ字読みをするため)

SMARTTALK・・・"amg" (「ア」、「メートル」、「グラム」と読むため)

以上の点から厳密な意味でのテストの公平性は保たれていない。しかし、評定者側には以上の点に関する事前情報は与えておらず、課題をやっていることに気がつく可能性は非常に低いので、実質的には比較可能であると考えられる。

このようにして生成したテキストを11種類の合成器に入力することによって合成音声を得た。尚、このような文字列の読み上げ方には合成器によって多様なクセがある。人間が一息に読み上げたように滑らかに早口で読む合成器もあれば、ひとつひとつをじれったいほど区切って読むものもある。これらの口調の統一は特に行わず、合成器が出すままに提示することとした。

4.1.2. 手続き

評定者は毎試行合成音によるアルファベット系列の読み上げを聞き取り、解答用紙に聞

き取った結果を書き取することを求められた。刺激の提示はラップトップパソコンによって制御した。被験者は各試行、準備が整ったら画面上をまずクリックすることによってその試行の課題音声聞くことができた。聴くのは各試行1回だけとした。聞き取れただけ書き取ったら、マウスクリックして次の試行へ進んだ。1文字ずつ聴いては順次書き取っていても、また全部を聴いてから記憶を辿って書き取っても、方略は評定者が最もやりやすいと思う方法をとって良いこととした。

刺激の提示順は、一連の11試行をブロック化し、1ブロック内にはすべての合成器による合成音が必ず1回ずつ提示されるという制約を除いてランダムとした。各評定者は各合成器、各文字数について2回ずつテストを実施した。従って評定者一人当たりの試行数は110試行であった。本試行に先立って11試行の練習を実施した。

4.1.3. 評定者

評定者は実験1, 2に参加した40名であった。

4.2. 結果と考察

回答が入力文字列に完全一致した割合(完全正答率)を文字数と合成器別に算出した結果を図12に示す。割合が高く、かつ入力文字数の増加に対して完全正答率の低下が激しくないほど聞き取りが容易であったことを示す。実験1, 2ではあまり高い評価を受けていないVisual Speech Creatorによる成績が最も良いことがこのグラフから分かる。報告者が聴く限りにおいてVisual Speech Creatorによる読み上げは最も話速が遅く、かつひとつひとつのアルファベットの区切りがはっきりとしたものである。おそらくこの合成器の音声に対してはひとつひとつを書き取っていても十分に間に合うはずで、そのために文字数が増加しても完全正答率の低下が起きにくいのであらうと考えられる。次に正答率の高いRealSpeakも同様の合成音をもたらしていた。報告者の主観的な印象ではアルファベットの読み上げ方はFine Speechが最も自然性が高く感じられた。その分話速は早く聞き流す分には問題がないが、一文字一文字を書き取るという場合には不利になっていたと考えられる。この合成器の場合、文字数が5から6へとひとつ増えたときに非常に大幅な正答率の現象が生じるという特徴を持っている。XIMERAはこの課題については中庸な結果を示している。

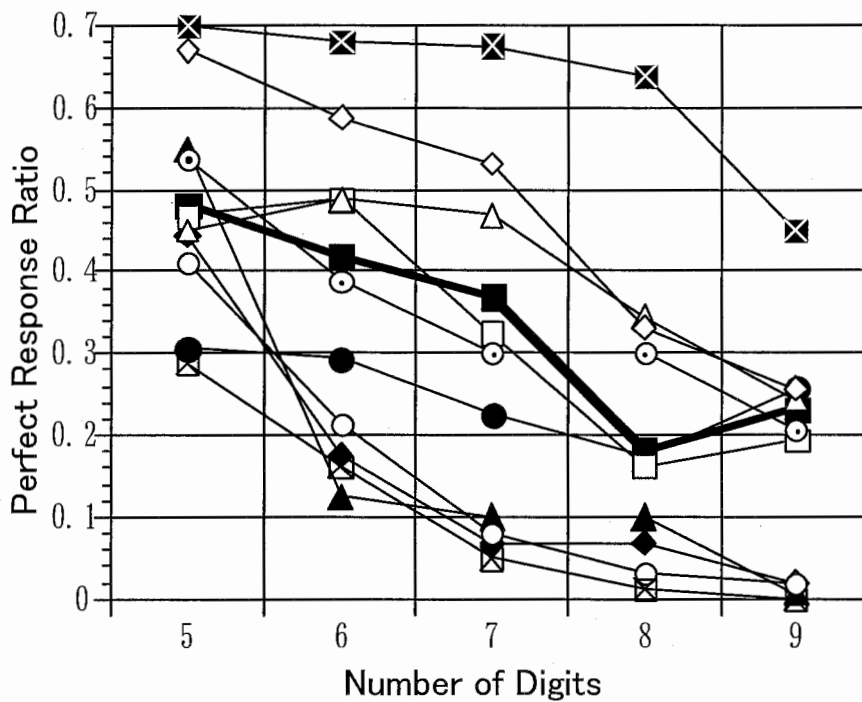
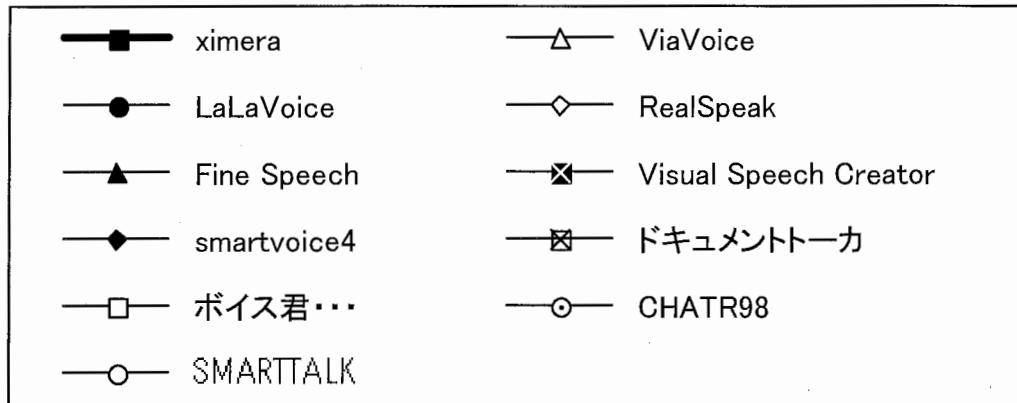


図 12 実験3において入力文字数の関数として表した完全一致の割合. 折れ線グラフの各線分は合成器の違いに対応する.

完全一致したものだけを正当と見なすという基準だけだと部分的に伝わった情報の程度については適正に現れない危険性がある. その点について補充する目的で, 回答中最長一致する文字列の数を数え, その刺激文字列数に対する割合を指標とする分析も施してみた. その結果が図 13 である. 完全一致正答率と比較した結果と本質的には変わらない結果を見なせる. 即ち, 完全一致基準を用いているとは言え, 相当数の母数をもとに算出した正答率は部分的な情報の伝達の側面についても表していると解釈できる.

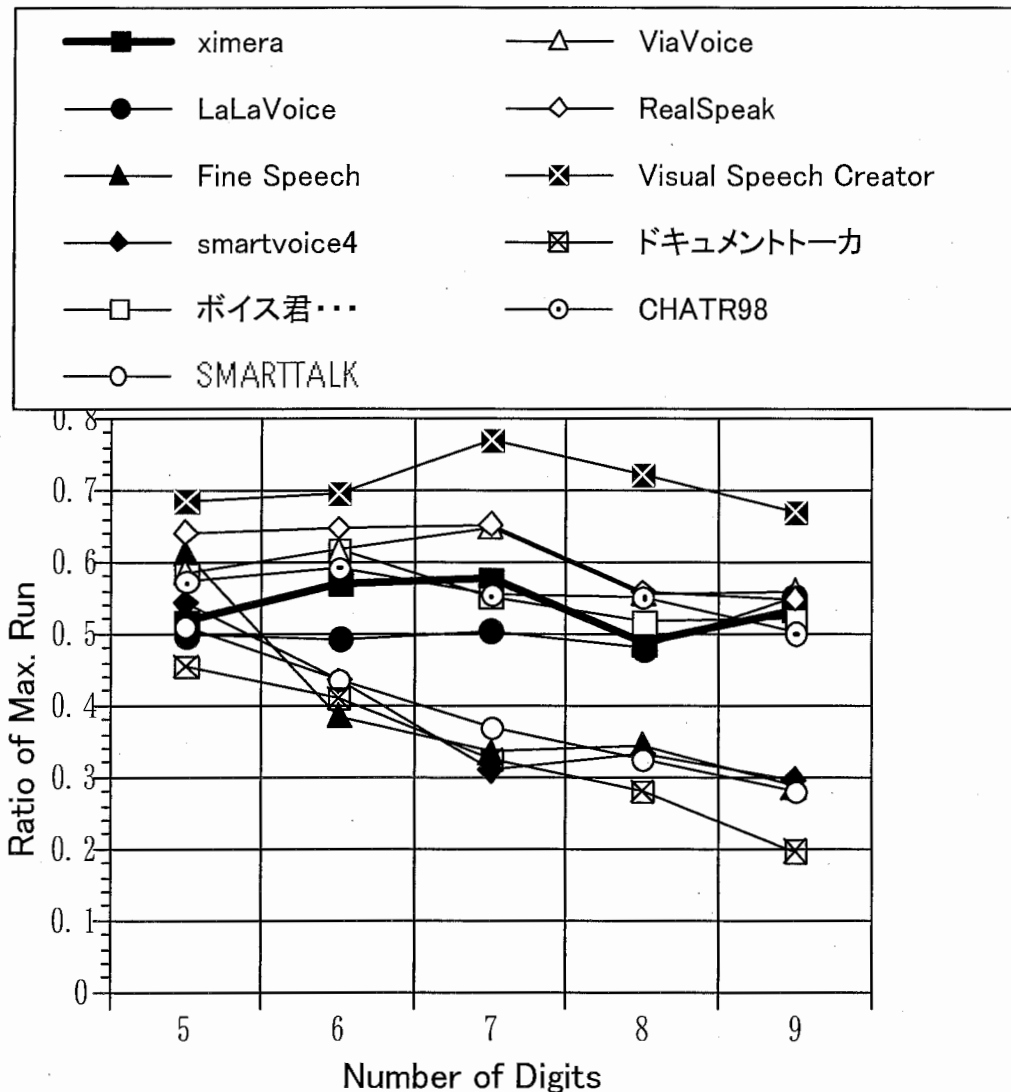


図 13 実験 3 において入力文字数の関数として表した最長一致の文字数の入力文字数への比。折れ線グラフの各線分は合成器の違いに対応する。

この実験は、実験 1, 2 とは異なる結果をもたらした。実験 3 での優位性を示した Visual Speech Creator の音質は実験 1 の結果を見る限りそれほど優秀であるとは考えにくく、また報告者自身が試聴した印象からも実験 3 の結果は合成音声の品質を測定しているとは言いがたい。XIMERA について実験 1, 2 では良好な評価結果を残してきた LaLaVoice での成績が実験 3 で伸びないこともこれを裏付けるものと考えられる。

実際に人間が口頭でアルファベットを伝える場合を考えると、本当に一文字一文字を正確に伝える場合はゆっくりと、しかも一文字をいうごとに相互に確認までとりながら行うことが多いはずである。よく起こるであろう状況では、まず最初に少し早口で滑らかに一連の文字列を発話し、その後からもう一度、今度は一文字一文字を区切って明瞭に読んで書きとどめさせるというモード使い分けが起こることが多いであろう。このようなモード

の使い分けに関する情報は現状の TTS ではしっかりと扱われていない。実際に XIMERA においてもアルファベットや数の読み上げについては特別な処理は行っていない。このような読み上げについては明らかに普通の文章とは異なるモードの使用を考えるべきである。

5. 全体考察

実験 1, 2 を通じて今回開発した XIMERA は現状で一般に手に入る TTS システムに対して明らかに優位性を示したと結論できる。LaLaVoice とは平均点において有意差が存在していただけでなく、評定のカテゴリの内訳を見ても最高点を付与される頻度が XIMERA の方が圧倒的に多いという質の違いを見せる。これは LaLaVoice が比較的小さな容量の素材を信号処理によって加工することによって合成を実現しているのに対して、XIMERA においてはいわゆる持続長変形や F0 変形などの信号処理は行わず、その代わりに大規模な波形コーパスを使用しているという基本構想の違いに起因する部分が多いと思われる。LaLaVoice の場合はほとんどの合成音について信号処理を受けたことにより「合成」されたことが人間の耳に露呈するのに対して、XIMERA については場合によっては自然音声との聞き分けがつかない程度にまで高品質な合成が提供可能だということであろう。

このような潜在能力がありながら、平均評定値自体にはそれほど大きな差が存在していないし、XIMERA の平均評定値も-3 から+3 の7段階評価にして+1 に達していない。+1 は「どちらかというが良い」というレベルである。平均値においてこのようなことになるのは、実験 1 の考察においても述べたように問題が生じる場合の劣化の程度については依然として大きいからと考えられる。

資料：音質評価時の辞書登録単語

1. 評価対象とした TTS システム

番号	製品名	バージョン	開発/発売元	API	記号	話者名
0	XIMERA		ATR	独自	ximera	
1	CHATR98		ATR	独自	catr98	
2	ボイス君のテキストスピーチ	V1.0.1.0	NTT データ	独自	voictg	
3	smartvoice4 xp		NEC	SAPI5	smrtvc	Female
4	SMARTTALK for WindowsV3	V3.01	沖電気	独自	smrttk	
5	ViaVoice for win standard	R10.0J	IBM	SAPI4	viavoi	日本語・標準 Hanako (女性)
6	LaLavoice2001		東芝	SAPI4	lalavo	お姉さん
7	ドキュメントトーカー Ver3.5	V3.5	クリエートシステム	SAPI4	dtalkr	花子
8	Visual Speech Creator		NTT-IT	独自	vsspch	
9	Fine Speech	V2.1	アニモ/富士通	SAPI5	fnspch	ケイコ
10	RealSpeakV2		スキャンソフト	独自	realsp	

2. 各アプリケーションの誤り単語数と登録単語数

番号	アプリケーション	初期状態で生じた誤り数	登録した単語数	単語登録語残った誤り数
0	ximera	6	6	0
1	catr98	16	16	0
2	voictg	14	14	0
3	smrtvc	6	6	0
4	smrttk	6	6	0
5	viavoi	13	0	13
6	lalavo	8	8	0
7	dtalkr	10	10	0
8	vsspch	11	11	0
9	fnspch	12	12	0
10	realsp	14	14	0

IBM 製品の ViaVoice は単語の登録ができなかったため、単語登録なしで評価実験を行った。

3. 辞書に登録した単語

アプリ名	単語	読み	修正前	修正後
ximera	洗濯物	せんたくもの	せんたくぶつ	せんたくもの
ximera	天然物	てんねんもの	てんねんぶつ	てんねんもの
ximera	逃避行	とうひこう	とうひぎょう	とうひこう
ximera	宇宙人	うちゅうじん	うちゅうひと	うちゅうじん
ximera	円買い	えんがい	えんがい	えんかい
ximera	豆球	まめきゅう	まめだま	まめきゅう

アプリ名	単語	読み	修正前	修正後
catr98	後手にまわって	ごてにまわって	うしろで	ごて
catr98	財務諸表が	ざいむしょひょうが	ざいむしょ、ひょうが	ざいむしょひょうが
catr98	円買い	えんがい	えんかい	えんがい
catr98	前日終値比	ぜんじつおわりねひ	ぜんじつおわりねふいりぴん	ぜんじつおわりねひ
catr98	係員	かかりいん	かかりーん	かかりいん
catr98	お止め	おやめ	おとめ	おやめ
catr98	本八幡	ほんやはた	ほんやわた	ほんやはた
catr98	3階	さんがい	さんかい	さんがい
catr98	明治通り	めいじどおり	あきはるとおり	めいじどおり
catr98	三国	みくに	さんごく	みくに
catr98	花恵	はなえ	はなめぐむ	はなえ
catr98	廣瀬橋	ひろせばし	ひろせばし	ひろせばし
catr98	初の単独	はつのたんどく	はじめのたんどく	はつのたんどく
catr98	空間	くうかん	あきま	くうかん
catr98	豆球	まめきゅう	まめたま	まめきゅう
catr98	寂しがり屋	さびしがりや	さびしがりおく	さびしがりや

アプリ名	単語	読み	修正前	修正後
voictg	その後	そのご	そのあと	そのご
voictg	元島清治	もとじまきよはる	もとじましおさむ	もとじまきよはる
voictg	最高潮位	さいこうちょうい	さいこうちょうくらい	さいこうちょうい
voictg	三国	みくに	さんごく	みくに
voictg	山田花恵	やまだはなえ	やまだばなえ	やまだはなえ
voictg	志田美保子	しだみほこ	こころざしたみほし	しだみほこ
voictg	止め下さい	やめください	とめください	やめください
voictg	生息	せいそく	しょうそく	せいそく
voictg	天然物	てんねんもの	てんねんぶつ	てんねんもの
voictg	土日	どにち	つちび	どにち
voictg	豆球	まめきゅう	まめたま	まめきゅう
voictg	日和	ひより	ひわ	ひより
voictg	本八幡	ほんやはた	ぼんはちまん	ほんやはた
voictg	廣瀬橋	ひろせばし	こうせばし	ひろせばし

アプリ名	単語	読み	修正前	修正後
smrtvc	一日	いちにち	ついたち	いちにち
smrtvc	天然物	てんねんもの	てんねんぶつ	てんねんもの
smrtvc	三国	みくに	さんごく	みくに
smrtvc	西	にし	さい	にし
smrtvc	最高潮位	さいこーちょーい	さいこうちょうくらい	さいこうちょうい
smrtvc	火口	かこー	ほくち	かこう

アプリ名	単語	読み	修正前	修正後
smrttk	お止めください	おやめください	おとめください	おやめください
smrttk	上下線	じょうげせん	じょう、かせん	じょうげせん
smrttk	近く	ちかく	きんく	ちかく
smrttk	天然物	てんねんもの	てんねんぶつ	てんねんもの
smrttk	当館	とうかん	とうやかた	とうかん
smrttk	豆球	まめきゅう	まめたま	まめきゅう

アプリ名	単語	読み	修正前	修正後
lalavo	お止め	おやめ	おとめ	おやめ
lalavo	コスモ石油	こすもせきゆ	こすもせきあぶら	こすもせきゆ
lalavo	一番館	いちばんかん	いちばんやかた	いちばんかん
lalavo	映画祭	えいがさい	えいがまつり	えいがさい
lalavo	花恵	はなえ	はなさと	はなえ
lalavo	国交相	こっこうしょう	こっこうそう	こっこうしょう
lalavo	天然物	てんねんもの	てんねんぶつ	てんねんもの
lalavo	後ほど	のちほど	あとほど	のちほど

アプリ名	単語	読み	修正前	修正後
dtalkr	お止め	おやめ	おとめ	おやめ
dtalkr	為替市場	かわせしじょう	がわせしじょう	かわせしじょう
dtalkr	宇宙人	うちゅうじん	うちゅうひと	うちゅうじん
dtalkr	角	かど	かく	かど
dtalkr	空いてる	あいてる	すいてる	あいてる
dtalkr	三国	みくに	さんごく	みくに
dtalkr	都美術館	とびじゅつかん	とびじゅつやかた	とびじゅつかん
dtalkr	博物館展	はくぶつかんてん	はくぶつかんあきら	はくぶつかんてん
dtalkr	後手	ごて	こて	ごて
dtalkr	本八幡	ほんやはた	ほんはちまん	ほんやはた

アプリ名	単語	読み	修正前	修正後
vsspch	8分	はっぶん	はちぶ	はっぶん
vsspch	お止め	おやめ	おとめ	おやめ
vsspch	悪口	わるぐち	あっこう	わるぐち
vsspch	言って	いって	ごとって	いって
vsspch	後ほど	のちほど	あとほど	のちほど
vsspch	最高潮位	さいこうちょうい	さいこうちょうぐらい	さいこうちょうい
vsspch	三国	みくに	さんごく	みくに
vsspch	山田花恵	やまだはなえ	やまだはなめぐみ	やまだはなえ
vsspch	寂しがり屋	さびしがりや	さびしがりおく	さびしがりや
vsspch	天然物	てんねんもの	てんねんぶつ	てんねんもの
vsspch	廣瀬橋	ひろせばし	ひろせはし	ひろせばし

アプリ名	単語	読み	修正前	修正後
fnsrch	お買得	おかいどく	おばいとく	おかいどく
fnsrch	お止め	おやめ	おとめ	おやめ
fnsrch	角	かど	かく	かど
fnsrch	最高潮位	さいこうちょうい	さいこうちょうぐらい	さいこうちょうい
fnsrch	寂し	さびし	じゃくし	さびし
fnsrch	人脈作り	じんみゃくづくり	じんみゃくつくり	じんみゃくづくり
fnsrch	その数	そのかず	そのすう	そのかず
fnsrch	天然物	てんねんもの	てんねんぶつ	てんねんもの
fnsrch	都美術館	とびじゅつかん	みやこびじゅつかん	とびじゅつかん
fnsrch	廣瀬橋	ひろせばし	ひろせきょう	ひろせばし
fnsrch	豆球	まめきゅう	まめたま	まめきゅう
fnsrch	三国	みくに	さんごく	みくに

アプリ名	単語	読み	修正前	修正後
realsp	お止め	おやめ	おとめ	おやめ
realsp	一日	いちにち	ついたち	いちにち
realsp	宇宙人	うちゅうじん	うちゅうひと	うちゅうじん
realsp	故障車	こしょうしゃ	こしょうぐるま	こしょうしゃ
realsp	後	あと	のち	あと
realsp	寂しがり屋	さびしがりや	さびしがりおく	さびしがりや
realsp	人脈作り	じんみゃくづくり	じんみゃくづっくり	じんみゃくづくり
realsp	盛りだくさん	もりだくさん	さかりだくさん	もりだくさん
realsp	銭	せん	ぜに	せん
realsp	銭安	せんやす	せんあん	せんやす
realsp	前総裁	ぜんそうさい	まえそうさい	ぜんそうさい
realsp	逃避行	とうひこう	とうひゆき	とうひこう
realsp	頭	とう	あたま	とう
realsp	廣瀬橋	ひろせばし	ひろせはし	ひろせばし

XIMERA (Ver1.1)

ベンチマーク報告書

(リソース編)

2004年6月

株式会社 国際電気通信基礎技術研究所

Advanced Telecommunications Research Institute International

目次

1.	はじめに	1
2.	測定対象の TTS システム	1
3.	測定項目	1
4.	テストデータ	1
5.	測定条件	1
5.1.	ハードウェアおよび OS	1
5.2.	音声データのサンプリング周波数	1
5.3.	データベース規模	2
6.	測定結果	3
6.1.	XIMERA 測定結果	3

1. はじめに

XIMERA の計算速度、消費メモリー量、所要ハードディスク容量に関して、音声データベース規模をパラメータとして測定を行いました。

2. 測定対象の TTS システム

XIMERA は、バージョン 1.1 を測定対象としました。

3. 測定項目

次の 3 項目について測定を行いました。

- (1) 音声合成処理に要した計算時間(モジュール別に測定)
- (2) 実行時に使用したメモリー量
- (3) インストールに必要なハードディスク容量

4. テストデータ

テストセットの文は、音素バランス文と長文の 2 種類を使用しました。

音素バランス文は、ATR B セット (音素バランス文) 中の J セット 53 文です。平均文長は、23.1 モーラ/文です。長文は、新聞テキストコーパスより採取した文セットであり、平均的な文長は、212 モーラ/文です。

5. 測定条件

5.1. ハードウェアおよび OS

- ・測定条件 1W・・・ハイエンド PC を Windows で使用

OS : WindowsXP CPU : Pentium4 2.8GHz、Chipset : Intel850MV、Memory : 2GByte、HardDisk : RAID0

- ・測定条件 1L・・・ハイエンド PC を Linux で使用

OS : Redhat7.2、CPU : Pentium4 2.4GHz、Chipset : Intel845GE、Memory : 2GByte

5.2. 音声データのサンプリング周波数

音声データのサンプリング周波数は、16kHz と 48kHz の 2 通りで測定を行いました。次章以

降では、それぞれの条件を 16k および 48k と表記します。

5.3. データベース規模

音声データベースの規模は、0.5 時間、47 時間の 2 通りについて測定を行いました。

6. 測定結果

6.1. XIMERA 測定結果

表 1. 測定条件 1W(ハイエンド PC を Windows で使用)での計算時間

テスト項目	データベースサイズ	
	0.5h	47h
16k		
1 文の平均合成時間	0.761	2.675
(モーラ/秒換算)	30.3	8.6
RTF	0.265	0.928
テキスト処理時間	0.109	0.103
韻律予測処理時間	0.131	0.133
素片選択処理時間	0.228	1.744
素片接続処理時間	0.292	0.696
48k		
1 文の平均合成時間	2.031	4.024
(モーラ/秒換算)	11.3	5.7
RTF	0.706	1.403
テキスト処理時間	0.109	0.108
韻律予測処理時間	0.134	0.130
素片選択処理時間	0.220	1.743
素片接続処理時間	1.568	2.042
16k ※1		
1 文の平均合成時間		2.226
(モーラ/秒換算)		10.3
RTF		0.776
テキスト処理時間		0.119
韻律予測処理時間		0.134
素片選択処理時間		1.760
素片接続処理時間		0.212

※1 音声素片の使用頻度情報にもとづく音声ファイルの先読みあり。

※ 単位は秒。

※ テストセットは、ATR 音素バランス文の J セット。

表 2. 測定条件 1L(ハイエンド PC を Linux で使用)での計算時間

テスト項目	データベースサイズ	
	0.5h	47h
16k		
1文の平均合成時間	0.852	3.167
(モーラ/秒換算)	27.0	7.3
RTF	0.296	1.104
テキスト処理時間	0.072	0.076
韻律予測処理時間	0.167	0.168
素片選択処理時間	0.348	2.486
素片接続処理時間	0.265	0.437
48k		
1文の平均合成時間	2.898	5.079
(モーラ/秒換算)	7.9	4.5
RTF	1.008	1.770
テキスト処理時間	0.075	0.076
韻律予測処理時間	0.168	0.171
素片選択処理時間	0.352	2.490
素片接続処理時間	2.304	2.342
16k ※1		
1文の平均合成時間		3.021
(モーラ/秒換算)		7.6
RTF		1.054
テキスト処理時間		0.089
韻律予測処理時間		0.168
素片選択処理時間		2.500
素片接続処理時間		0.265

※1 音声素片の使用頻度情報にもとづく音声ファイルの先読みあり。

※ 単位は秒。

※ テストセットは、ATR 音素バランス文のJセット。

表 3. 長文テストセットに対する測定条件 1L(ハイエンド PC を Linux で使用) での計算時間

テスト項目	測定条件	
	47h (16k)	47h (16k) ※1
1 文の平均合成時間	28.316	27.171
(モーラ/秒換算)	7.5	7.8
RTF	0.977	0.936
テキスト処理時間	0.278	0.291
韻律予測処理時間	1.640	1.651
素片選択処理時間	23.368	23.380
素片接続処理時間	3.030	1.849

※1 音声素片の使用頻度情報にもとづく音声ファイルの先読みあり。

表 4. 使用ディスク容量及び合成時メモリーの使用量

テスト項目	データベースサイズ	
	0.5h	47h
XIMERA ディスク使用量 (16k)	0.3	10.7
XIMERA ディスク使用量 (48k)	0.4	28.2
XIMERA メモリー使用量 (16k, ATR503)	67	391
XIMERA メモリー使用量 (16, 長文)		596

※ 測定条件は 1W(ハイエンド PC を Windows で使用)。

※ ディスク使用量の単位は GByte, メモリー使用量の単位は MByte。