TR-IT-0056

音声言語データベースの構成

ATR Integrated Speech and Language Database

浦谷 則好　竹沢 寿幸　松尾 秀彦*　森田 千帆*

Noriyoshi URATANI　Toshiyuki TAKEZAWA　Hidehiko MATSUO　Chiho MORITA

１９９４年５月

概　要

ＡＴＲで構築中の音声言語データベースについて、その特徴、ファイルの構成、ファイル名の命名法、属性情報の中身などについて解説する。

ＡＴＲ音声翻訳通信研究所

ATR Interpreting Telecommunications Research Laboratories

---

*株式会社 東洋情報システム

TOYO Information System CO., LTD.

目　次

１．はじめに　・・・　1

２．音声言語データベースの特徴　・・・　2

３．ファイルの構成　・・・　5

４．ファイル名の命名法　・・・14

５．会話属性情報　・・・17

６．おわりに　・・・25

参考文献　・・・26

１．はじめに

当所の前身であるＡＴＲ自動翻訳電話研究所は音声データベースと言語データベースをそれぞれ別個に構築していた[1][2][3]。音声データベースは主として音声認識のための音素モデル構築に利用する目的で収集されたものであり、そのほとんどは単語発声や文節発声のものである。そのため、文発声を対象としたい研究者にとっては不満なものであった。一方、言語データベースは自然発話を対象としていたため、近未来の音声認識システムを目標に置く研究者にとっては手に余るものであった。そもそも、音声言語を対象としたい研究者にとって音声データと言語データの関連のとれたデータベースの不在は研究進展のためのネックとなっていた。そこで、ＡＴＲ音声翻訳通信研究所は音声データと言語データの関連のとれたデータベース、すなわち音声言語データベース、を構築することを決めた[4]。

本報告書は、ＡＴＲが構築中の音声言語データベースの概要を解説したものである。内容は、２．で音声言語データベースの特徴について触れ、３．で音声言語データベースがどのようなファイルによって構成されているかをを述べ、４．で各ファイルの名称がどのような決まりの下に付けられているかを示し、５．会話の属性情報としてどのようなものが付与されているかを説明する。

2．音声言語データベースの特徴

音声言語データベースの最大の特徴は前述したように、音声データと言語データが関連していることである。音声データとしての特徴としては文発声（正確には人が自然に発声するときのまとまりで、複数文を含むこともある）であること、朗読ではなく自然な発声（に近い）のデータであることが挙げられる。言語データとして見たときには、ＡＴＲ自動翻訳電話研究所が作成した旧言語データベースと同様、バイリンガルのデータベースであり、

・日英の対応データが存在する

・基礎的な言語解析情報が付与される

という特徴はこのデータベースにも継承されている。但し、日本語の書き起こしの際の表記や形態素解析情報に関しては大幅な変更を行なっている[5][6]。日本語構文情報や英語形態素情報に関しても変更する予定であるが、この報告書作成の時点においてまだ細部が決まっていないので、これらについては別途報告する予定である。

対象タスクは「旅行に関する会話」としている。これは、旧言語データベースとのつながりを重視したからであり、かつ旧言語データベースの対象タスクの１つであった「国際会議に関する問い合わせ」より広い応用を狙ったからである。

データの収集方法も旧言語データベースを踏襲していて模擬実験対話によって行なっている。但し、品質の良い音声データを確保するために以下の条件を付けている。

・収録はスタジオなどの環境ノイズが小さいところで行なう

・マイクは予めＡＴＲが指定する　（既知の特性のマイクを使用）

・記録にはＤＡＴを用いる

さらに、旧言語データベースのように自由な会話を前提とするのではなく、近未来に音声翻訳システムが使用される状況をイメージして、日本語話者と英語話者による、通訳者（日英方向と英日方向別々に計２名）を介する会話を前提としていることのほかに、以下のような制限も付けている。

・１発話は１０秒以内であること

・割り込み発話（相槌も含む）がないこと

・極端に長い言い淀みや言い直しがないこと

旧言語データベースのように「全ての会話に均一の情報を付与する」ことを基本としていないことに注意されたい。むしろ、会話に対する情報の付与は「利用度の大きい（と期待される）もの」ほど多くすることを予定している。

データの作成工程を示すと図１および図２のようになる。これらのデータがどのように管理されているかは３．以降でふれる。

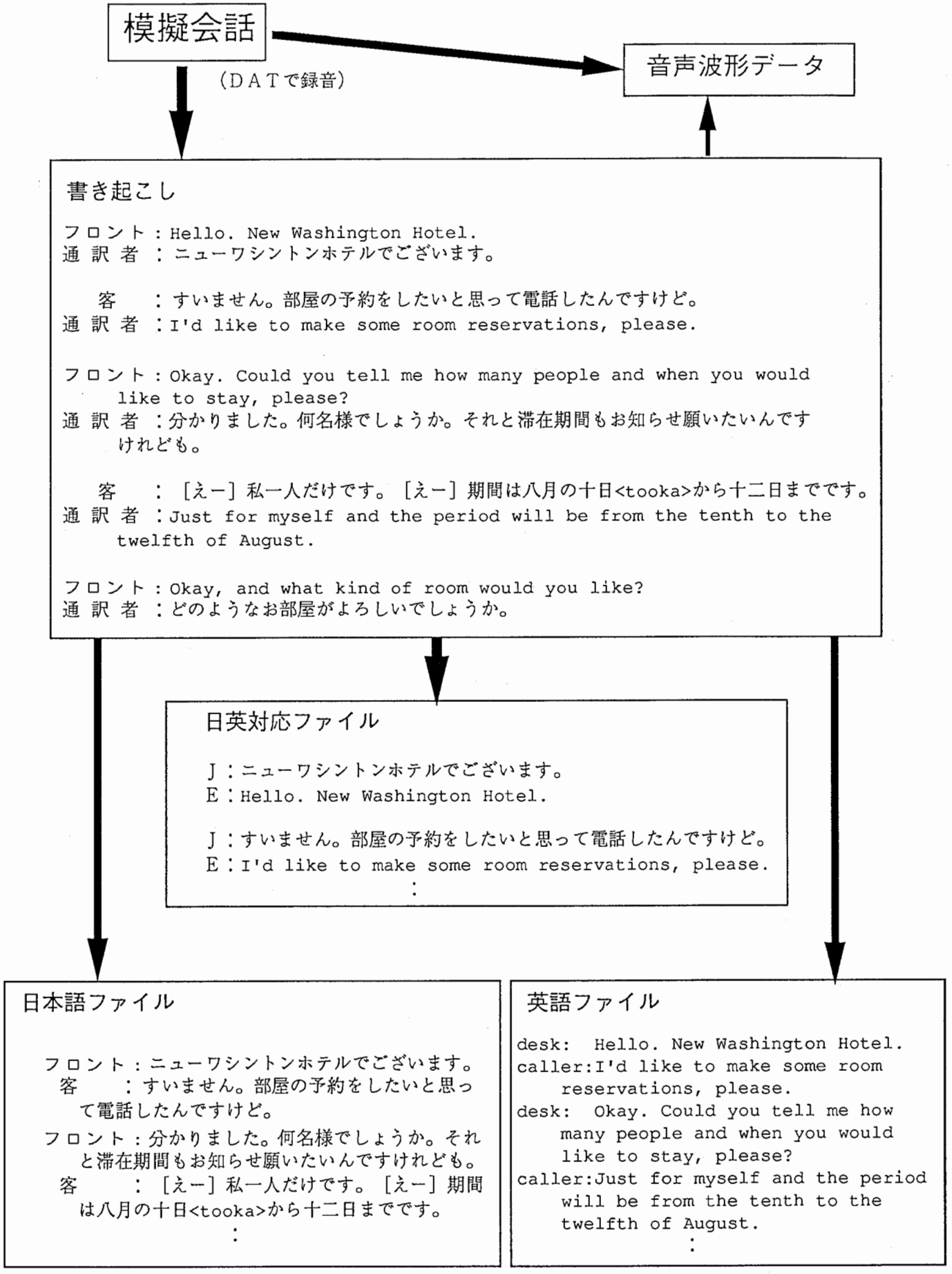

図１　データ作成の手順（１）

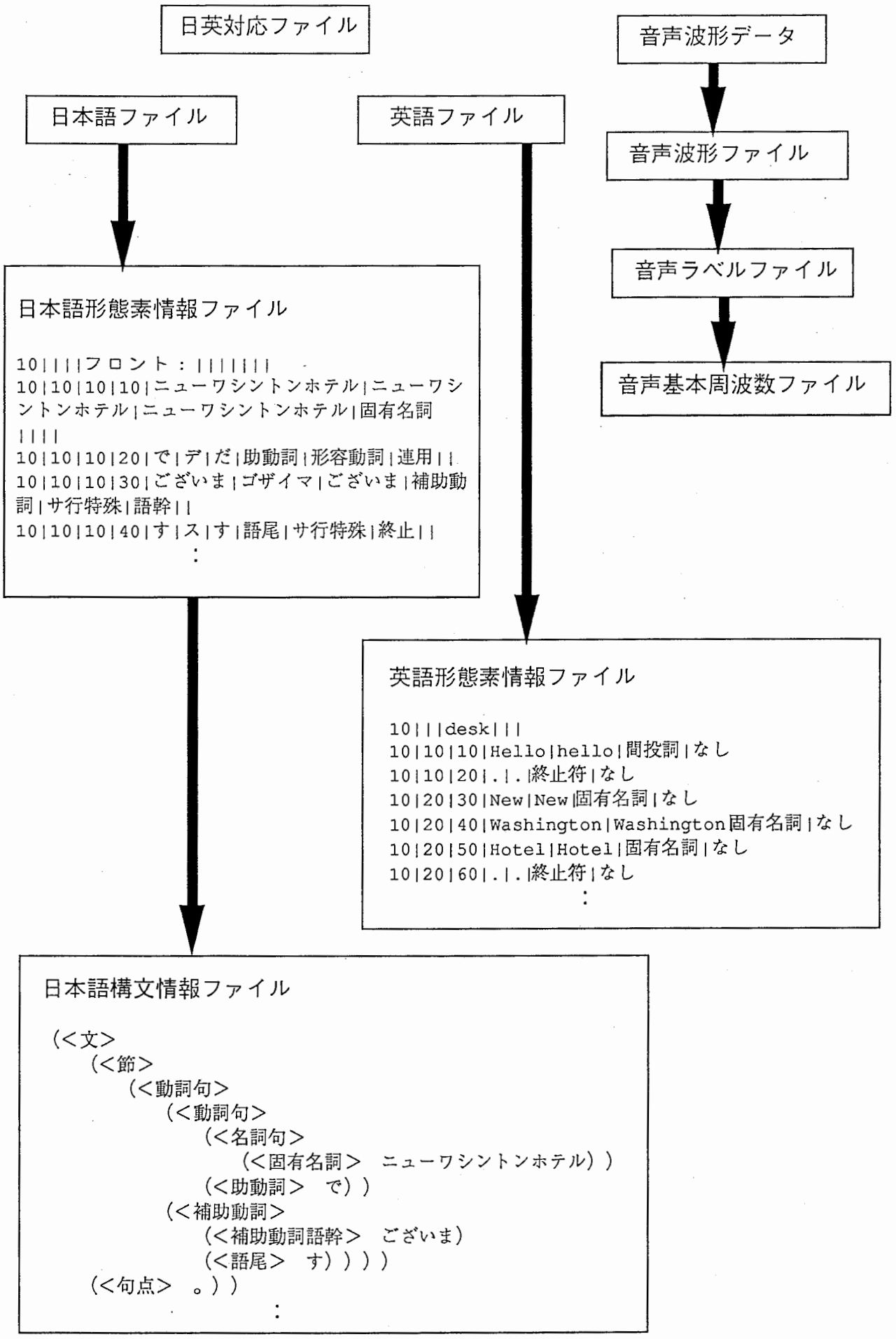

図２　データ作成の手順（２）

3．ファイルの構成

データの可視性に重点をおいて、データベースはＵＮＩＸのファイル・システムをそのまま利用してファイルを管理することとしている。言語データは基本的に１会話毎にファイル化され、データの種類毎にディレクトリに格納される。一方、音声データの方はデータ量が多いため１発話毎にファイル化され、１会話毎にディレクトリに格納される。この会話毎のディレクトリはデータの種類毎に設定された上位のディレクトリに格納される。したがって、音声言語データベースにおけるディレクトリ構成は

```
$SLDB[*1]/ENV/                  ;会話属性
$SLDB/LNG/                      ;言語データ
            JTEXT               ;    日本語会話文
            ETEXT               ;    英語会話文
            JESYM               ;    日英対応
            JMOR                ;    日本語形態素
            EMOR                ;    英語形態素
            JTREE               ;    日本語構文木
$SLDB/SPH/                      ;音声データ
            WAV/                ;    音声波形
                会話ID          ;        会話単位毎の音声波形
            LBL/                ;    音声ラベル
                会話ID          ;        会話単位毎の音声ラベル
            PIT/                ;    音声基本周波数
                会話ID          ;        会話単位毎の音声基本周波数
```

のようになっており、これを図示すると図３のようになる。同一会話に属する各種のデータはファイル名によって関係付けられている。（この命名法については４．でふれる。）

以下で各ファイルの形式についてふれる。

[*1] １９９４年５月現在、$SLDB は /DB/SLDB として全ての研究員のUNIX WS から利用することが可能である。

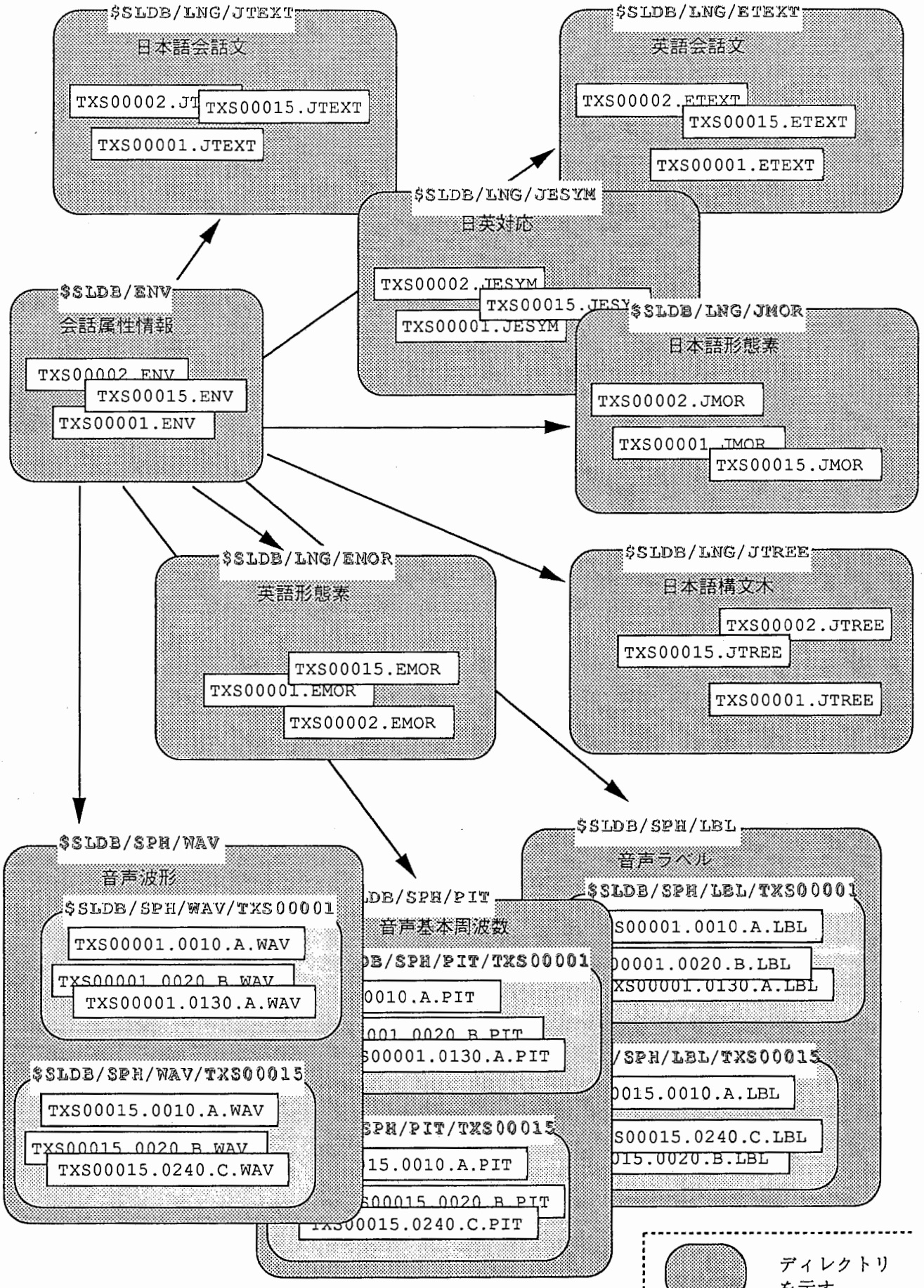

図３　データ管理のイメージ

3．1　各データの形式

3．1．1　会話属性ファイル

会話属性情報には言語データ、音声データが収集された際の諸条件を記載する。

●ファイル名　　会話ID.ENV

●ファイル形式　　5．会話属性情報　を参照

3．1．2　言語データ

3．1．2．1　日本語会話文

●ファイル名　会話ID.JTEXT

●ファイル形式

<File>　:= <日本語発話>+

<日本語発話>　:= 日本語発話マーク[*1]　発声[*2]　'NL' <同一発話内の２発声目以降>

<同一発話内の２発声目以降>*　:= ４全角ｽﾍﾟｰｽ　発声

<発声>　:= <文>+

('NL'：　改行を示す)

●例

通訳者：もしもし。

担当者：はい、京都ツーリストでございます。

通訳者：［あの］私たち、アメリカから来たグループなんですが、実は日本置伝統的なお料理を京都情緒の中で味わってみたいのですけれど。

：

担当者：いいえ、そのご予算でもすてきなお弁当を召し上がっていただけますよ。

通訳者：分かりました。では、よろしくお願いします。

[*1] 8バイト使用し、会話者の役割を示すマークを付ける。
"申込者：", "担当者：", "通訳者："などがある。

[*2] 発話者が長いポーズを入れるまでの発声のまとまりのこと。
通訳者を介する会話では、通訳を依頼する文のまとまりを指す。

3．1．2．2　英語会話文

●ファイル名　会話ID.TEXT

●ファイル形式

<File>　:= <英語発話>+

<英語発話>　:= 英語発話マーク[*1]　発声　'NL' <同一発話内の２発声目以降>

<同一発話内の２発声目以降>*　:= 13半角ｽﾍﾟｰｽ　発声

<発声>　:= <文>+

●例

customer:　Hello.

interpreter:　Kyoto Tourist.

customer:　[uh] We're a group of American tourists and we'd like to try some [uh] traditional Japanese food in a traditional Kyoto environment.

...

interpreter:　Yes, within your budget you mentioned we can serve you very nice Japanese style box lunch.

customer:　OK. I'll do that. Thank you very much.

3．1．2．3　日本語形態素情報

日本語会話に対して文や文節の情報、及び形態素情報を付加する。

●ファイル名　会話ID.JMOR

●ファイル形式

<File>　:= <形態素>+

<形態素>　:= 発話ID　'|' 発声ID　'|' 文節ID　'|' 形態素ID　'|' 表記形　'|' 読み　'|'標準形　'|' 品詞　'|' 活用１　'|' 活用２　'|' コメント　'|'

[*1] 13バイト使用し、会話者の役割を示すマークを付ける。
"clerk","customer","front desk","interpreter"などがある。

●例

10¦¦¦¦通訳者：¦¦¦¦¦¦

10¦10¦10¦10¦もしもし¦もしもし¦もしもし¦感動詞¦¦¦¦

10¦10¦10¦20¦。¦¦。¦記号¦¦¦¦

：

10¦90¦230¦390¦すてき¦すてき¦すてき¦形容名詞¦¦¦¦

10¦90¦230¦400¦な¦な¦だ¦助動詞¦連体形¦¦¦

10¦90¦240¦410¦お¦お¦お¦接頭辞¦¦¦¦

10¦90¦240¦420¦弁当¦べんとう¦弁当¦普通名詞¦¦¦¦

3．1．2．4　英語形態素情報

英語会話に対して文や形態素の情報を付加する。

●ファイル名　会話ID.EMOR

●ファイル形式

<File>　:= <形態素>+

<形態素>　:= 発話ID '¦' 発声ID '¦' 形態素ID '¦'表記形 '¦' 標準形 '¦'
品詞 '¦' 活用１ '¦' 活用２ '¦' コメント '¦'

●例

10¦¦¦¦customer:¦¦¦¦

10¦10¦10¦Hello¦hello¦間投詞¦なし¦なし¦¦

10¦10¦20¦.¦.¦終止符¦なし¦なし¦¦

…

10¦90¦330¦can¦can¦助動詞¦現在形¦なし¦¦

10¦90¦340¦serve¦serve¦動詞¦原形¦なし¦¦

10¦90¦350¦you¦you¦代名詞¦目的格¦複数¦¦

10¦90¦360¦very¦very¦副詞¦なし¦なし¦¦

…

3．1．2．5　日英対応情報

●ファイル名　会話ID.JESYM

●ファイル形式

<File>　:= <日英対応>+

<日英対応>　:= <日本語文>+　'|' <発声番号> 'NL' <英語文>+　'|' <発声番号>

<日本語文>　:= <日本語表記>+ :

<英語文>　:= <英語表記>+ :

<日本語表記>　:= 表記形 'SP'

<英語表記>　:= 表記形 'SP'

<発声番号>　:= 10,20 .....

('SP'：１つ以上の半角スペース)

●例

もしもし、通訳電話国際会議事務局ですか？¦10

Hello , is this the Secretariat for the International Conference of Interpreting Telephony ? ¦ 10

わかりました。¦それでは、こちらからお送りしますので、お名前とご住所を...

Okay , then I will send them to you . : Would you please give me your name and address ? ¦ 30

3．1．2．6　日本語構文情報

●ファイル名　会話ID.JTREE

●ファイル形式

<File>　:= <日本語構文情報>+

<日本語構文情報>　:= ";;(" インデックス 'SP' "形態素後処理前 見出し"
'SP' 文 ")" 'NL'
";;(" インデックス 'SP' "形態素後処理前 品詞 " <品詞>+ ")" 'NL'
'SP' 文 ')' 'NL' '発話ID' '|'
";;(" インデックス 'SP' "形態素後処理後 見出し" 'SP' 文 ")" 'NL'
";;(" インデックス 'SP' "形態素後処理前 品詞 " <品詞>+ ")" 'NL'
";;(" インデックス 'SP' "形態素後処理前 品詞 " <品詞>+ ")" 'NL'
"(" インデックス 'NL' <TREE構造> ")" 'NL'

<インデックス>　:= '会話ＩＤ' "-" '発話ＩＤ' "-" '枝番号'

<TREE構造>　:= "(" '中間シンボル名' '終端記号' ")" 'NL'

<TREE構造>　:= "(" '中間シンボル名' <TREE構造> ")" 'NL'

セミコロン（';'）から改行までの文字列はコメントとみなす。

したがって、上記の形態素列の情報もコメントの一部であるが、必ず<TREE構造>の直前に付与することとする。

●例

;;(TS33001-800-1 形態素後処理前 見出しいいえ、その予算でもすてきなお弁当を召し上がっていただけますよ。)

;;(TS33001-800-1 形態素後処理前 品詞 感動詞 記号 連体詞 普通名詞 格助詞 係助詞 形容名詞 助動詞 接頭辞 普通名詞 格助詞 本動詞 語尾 助動詞 語尾 助動詞 語尾 終助詞 記号)

;;(TS33001-800-1 形態素後処理後 見出しいいえ、その予算でもすてきなお弁当を召し上がっていただけますよ。)

;;(TS33001-800-1 形態素後処理後 品詞 感動詞 読点 連体詞 普通名詞 格助詞 係助詞 形容名詞 助動詞 接頭辞 普通名詞 格助詞 本動詞 語尾 助動詞語幹 語尾 助動詞語幹 語尾 終助詞 句点)

```
(TS33001-800-1
(<文>
  (<節>
    (<感動詞>
      (<感動詞>　いいえ)
      (<読点>　、))
    (<節>
      (<動詞句>
        (<動詞句>
          (<動詞句>
            (<動詞>
              (<後置詞句>
                (<後置詞句>
                  (<名詞句>
                    (<連体詞> その)
                    (<名詞句>
                      (<普通名詞>　予算)))
                  (<格助詞　で))
                (<係助詞> も))
```

```
        (<動詞>
          (<後置詞句>
            (<名詞句>
              (<連体修飾節>
                (<動詞句>
                  (<形容名詞> すてき)
                  (<助動詞> な)))
                (<普通名詞>
                  (<接頭辞> お)
                  (<普通名詞> 弁当)))
              (<格助詞 を))
            (<動詞>
              (<本動詞> 召し上が)
              (<語尾> っ ))))
          (<助動詞>
            (<助動詞語幹> ていただ)
            (<語尾> け)))
        (<助動詞>
          (<助動詞語幹> ま)
          (<語尾> す)))
      (<終助詞> よ))))
  (<句点> 。))
)
```

３．１．３　音声データ

音声波形データは１会話の発話を１ファイルに格納すると膨大な量になるので、１発話ごとに１ファイルとして格納している。波形データとの対応を容易にするために他の音声データファイルも１発話１ファイルとしている。以下で述べるデータのほか、韻律情報に関するデータもできれば作成したいと考えているが、データ形式等が未定なためこの報告書ではふれない。

３．１．３．１　音声波形ファイル

●ファイル名　　会話ＩＤ．発声ＩＤ．発声者ＩＤ．ＷＡＶ

●ファイル形式　　バイナリーデータ（short integer:16 bit）

３．１．３．２　音声ラベルファイル

●ファイル名　　会話ＩＤ．発声ＩＤ．発声者ＩＤ．ＬＢＬ

●ファイル形式　　アスキーデータで

１．リスト
２．実発声（イベント）
３．異音化層
４．融合化層
５．母音中心層
６．コメント層

の６層からなる。（ただし、３，４，６は省略されることが多い）

３．１．３．３　音声基本周波数ファイル

●ファイル名　　会話ＩＤ．発声ＩＤ．発声者ＩＤ．ＰＩＴ

●ファイル形式　　バイナリーデータ（short integer:16 bit）
（5msec毎に抽出されたデータ）

4　ファイル名の命名法

音声言語データベースに関連する各々のデータの関連性を明確にし、活用しやすくしていくためにデータのファイル名の統一を図らなくてはならない。ここでは規則的なファイル名の命名法を定義する。

かつファイル名はすべて１ｂｙｔｅコードで記述し、２ｂｙｔｅコードは使用しないものとする。

4.1　各データ毎のファイル名のフォーマット

ファイルの属性	ファイル名
音声波形（４８ kHz）	会話ＩＤ.発声ＩＤ.話者ＩＤ.WAV.48K
音声波形（１２ kHz）	会話ＩＤ.発声ＩＤ.話者ＩＤ.WAV
音声ラベル	会話ＩＤ.発声ＩＤ.話者ＩＤ.LBL
音声基本周波数	会話ＩＤ.発声ＩＤ.話者ＩＤ.PIT
会話属性情報	会話ＩＤ.ENV
日本語会話文	会話ＩＤ.JTEXT
英語会話文	会話ＩＤ.ETEXT
日本語形態素解析	会話ＩＤ.JMOR
英語形態素解析	会話ＩＤ.EMOR
日本語構文解析	会話ＩＤ.JTREE
日英対応	会話ＩＤ.JESYM

4.2　会話ＩＤ

会話ＩＤは８桁（固定長）からなり、音声言語データベースの全てのデータに共通して付ける。

会話ＩＤのフィールドの設定は以下のようになっている。

表 1: 会話ＩＤのテーブル

桁	内容
１文字目	ドメイン
２文字目	トピック
３文字目	発話形態
４～８文字目	番号ＩＤ

4.2.1 ドメイン

会話ＩＤの１文字目にあたり、それぞれ以下の表の略号を用いて記述する。

表 2: ドメイン一覧

ドメイン	略号	Domain
国際会議	I	International Conference
旅行	T	Travel
ホテルの予約	R	Reservation (Hotel)
学校生活	P	School Life
ふるさと通信	C	Communication
会議の日程調整	S	Appointment Scheduling

4.2.2 トピック

ドメインがＴ（旅行）の場合、以下のトピックの略号を入れる。
また、ドメインがＴ（旅行）以外の場合は一律大文字の「Ｘ」とする。

表 3: トピック一覧

略号	トピック
A	ホテルの部屋の予約（フロント係との対話）
B	ホテルの紹介（インフォメーションとの対話）
C	ホテルでのサービス・トラブル（フロント係との対話）
D	会議室利用の問合わせ・申込み
E	未定
F	未定
G	飛行機のフライトの予約・変更・解約
H	バス・列車の切符の問合せ
I	交通手段の問合せ
J	道案内
K	レンタカーの問合せ・申込み
L	未定
M	未定
N	未定
O	バスツアー・旅行パックの問合せ・申込み
P	未定
Q	未定
R	演劇・コンサートのチケットの予約・購入
S	レストランの予約
T	料理の注文
U	未定
V	未定
W	未定

4.2.3 発話形態

会話ＩＤの３文字目にあたり、それぞれ以下の表の略号を用いて記述する。

表 4: 発話形態一覧

発話様式／媒体	略号	Speaking Style / Media
自由発話	S	Spontaneous
演技	P	Play
朗読	R	Reading
キーボード	K	Keyboard

4.2.4 番号ＩＤ

会話ＩＤの４文字目から８文字目にあたり、５ｂｙｔｅの固定長で００００１から１ずつの増分で会話ＩＤに付与するものとする。

4.3 発声ＩＤ

発話とは話者が交替するまでの発声のまとまりをさす。

発声単位はここでは話者が通訳者に通訳を依頼する発声のまとまりのことをさす。この発声単位毎にＩＤが付与される。

１発声が２文以上からなることもある。発声ＩＤは４ｂｙｔｅの固定長で００１０から１０づつの増分で会話ＩＤの後に「.」（ドット）を伴う形で付与する。

この発声ＩＤは「１発声＝１ファイル」の音声情報関連ファイルに対してのみ付けるため、「１会話＝１ファイル」のデータに対しては省略することができる。

4.4 話者ＩＤ

このＩＤは発声ＩＤに「.」（ドット）を伴う形で付与される、発声ＩＤに付属のＩＤである。このＩＤは発声ＩＤの無いファイル（「１会話＝１ファイル」）に付ける事はできない。
このＩＤによって、発呼者、被呼者、通訳者を区別することができる。

表 5: 話者ＩＤの一覧

略号	話者
A	発呼者（電話をかけた方の話者）
B	被呼者（電話を受けた方の話者）
C	通訳者（日 → 英または両方向）
D	通訳者（英 → 日）

5　会話属性情報

会話の属性情報は会話収録時の情報を管理、利用していくために不可欠なものである。
この情報は１対話に対し１ファイル作成される。
以下に「ファイル名の命名法」、「フォーマット」、「各フィールド毎の記述方法」を示す。

5.1　ファイル名の命名法

この会話属性情報のファイル名の命名法は、14ページの「4. ファイル名の命名法」にしたがって、『会話ＩＤ.ENV』とする。

5.2　全体フォーマット

会話ＩＤ:
収録日時:
収録社名:
総会話時間:
有効会話時間:
ノイズレベル:
マイクロホン:
ＤＡＴの機種:
ミキサーの機種:
サンプリング周波数:
発話形態:
ドメイン:
トピック:
言語パターン:
発声方法:
対面・非対面:
発呼者:
被呼者:
通訳者１:
通訳者２:
翻訳者:
コメント:

※『:』は１ byte コードとする。

5.3　各フィールド毎の記述方法

5.3.1　会話ＩＤ

会話ＩＤは「4. ファイル名の命名法」に従って決定されたＩＤを 1 byte コードで記述する。
（8 byte 固定）

＜例＞ 会話ＩＤ:TXS00001

5.3.2　収録日時

１９○○年○○月○○日と記述する。
数字は 1 byte コードとし、月日は 2 桁に固定する。

＜例＞ 収録日時:1969 年 10 月 04 日

5.3.3　収録社名

音声の収録を行なった会社名を記入する。

＜例＞ 収録社名: ＡＴＲ音声翻訳通信研究所

5.3.4　総会話時間

総会話時間とは会話に要した時間で、時間切れなどで無効となった発声も含んだ時間である。
会話時間は 1 byte コードで、『分・秒』は 1 byte コードで記述する。
（2 byte 固定＋「分」＋ 2 byte 固定＋「秒」）

＜例＞ 総会話時間:78 分 01 秒

＜例＞ 総会話時間:00 分 23 秒

5.3.5　有効会話時間

有効会話時間とは、総会話時間から時間切れなどで無効となる発声を省いた会話の時間。
会話時間は 1 byte コードで、『分・秒』は 1 byte コードで記述する。
（2 byte 固定＋「分」＋ 2 byte 固定＋「秒」）
計測不可能な場合は 1 byte コードの「NIL」を入れておく。

＜例＞ 有効会話時間:72 分 01 秒

＜例＞ 有効会話時間:00 分 10 秒

5.3.6　ノイズレベル

以下の収録場所から該当する収録環境を選択し、実際に行なったノイズの計測結果を「○○dBA ＜収録環境＞」の形で記述する。
『○○dBA』はすべて1 byte コードで5 byte 固定とする。

無響音室
遮音室
オフィスルーム（上記の無響音室・遮音室以外の収録環境全般）

＜例＞ ノイズレベル:70dBA ＜オフィスルーム＞

＜例＞ ノイズレベル:09dBA ＜遮音室＞

5.3.7　マイクロホン

会話収録に使用したマイクロホンについて「型番＜製造社名＞」の形ですべて2 byte コードで記述する。

＜例＞ マイクロホン: ＭＵ－２Ｃ＜ｓａｎｋｅｎ＞

5.3.8　ＤＡＴの機種

会話収録に使用したＤＡＴについて「型番＜製造社名＞」の形ですべて2 byte コードで記述する。

＜例＞ ＤＡＴの機種: ＸＤ－５０５＜Ｖｉｃｔｏｒ＞

5.3.9　ミキサーの機種

会話収録に使用したミキサーについて「型番＜製造社名＞」の形ですべて2 byte コードで記述する。

＜例＞ ミキサーの機種: Ａ－８１２ＥＸ＜ＯＮＫＹＯ＞

5.3.10　サンプリング周波数（量子化１６ bit 固定）

サンプリングの周波数を1 byte コードで記述する。

＜例＞ サンプリング周波数:48kHz

5.3.11 発話形態

以下の発話形態の表にしたがって、発話形態の部分を記述する。

表 6: 発話形態一覧

発話形態	略号	内容
自由発話	S	台本無しで会話を行なった場合。
演技	P	役者などが台本を覚えて演技をした場合。
朗読	R	台本やテキストなどを読み上げた場合。
キーボード	K	キーボードを使って会話した場合。

＜例＞ 発話形態: 自由発話

5.3.12 ドメイン

以下のドメインの表にしたがって、ドメインの部分を記述する。

表 7: ドメイン一覧

ドメイン	略号	内容
国際会議	I	国際会議に関する問合せ
旅行	T	旅行に関する問合せ
ホテルの予約	R	ホテルの予約に関する会話
学校生活	P	学校生活に関する会話
ふるさと通信	C	ラジオ番組
会議の日程調整	S	会議の日程調整についての会話

＜例＞ ドメイン: 会議の日程調整

5.3.13 トピック

15ページのトピック一覧表にしたがってトピックを記述する。

表中のトピックにあてはまらない時及び、旅行以外のドメインで会話を行なっている場合は、一律 1 byte コードの「NIL」を入れておく。

＜例＞ トピック: チケットの予約

＜例＞ NIL

5.3.14　言語パターン

会話を行なった2者又は3者の発声した言語について記述する。
同一言語同士での会話は『発呼者の話した言語－被呼者の話した言語』の形で記述し、異なった言語同士で通訳者を介して行なった会話については『発呼者の話した言語－通訳者－被呼者の話した言語』のように記述する。

＜例＞ 言語パターン: 日本語－日本語

＜例＞ 言語パターン: 日本語－通訳者－英語

5.3.15　発声方法

会話収録時の話者の発声方法について記述する。
発声方法については以下を参考にする。

文発声
短文発声
長文発声
単語発声
文節発声
音節発声

＜例＞ 発声方法: 文発声

5.3.16　対面・非対面

話者同士が対面しているという設定で会話を行なっているか、非対面で会話を行なっている設定かを2 byteコードで記述する。
原則的にはこの欄は非対面である。

＜例＞ 対面・非対面: 非対面

＜例＞ 対面・非対面: 対面

5.3.17　発呼者

電話をかけた方の話者（発呼者の話者ＩＤ＜Ａ＞にあたる人物）の情報を記述する。
『話者の名前_ 性別_ 年齢_ 出身地_ 職業』のフォーマットで記述する。

注意点

- アンダーバー「_ 」は 2 byte コードで記述する。
- 話者の名前は 2 byte コードで記述する。
- 話者の姓と名の間には 2 byte コードの中点（・）を入れること。
- 話者の名前は「姓」、「・」、「名」の順番で記述する。
- 漢字と平仮名で表現できない時は 2 byte コードの英字表記にする。この場合も、「姓・名」と記述する。
- 性別は「女性」又は「男性」と記述する。
- 年齢は 1 byte コードで 2 文字で記述する。（2 byte 固定）
- 出身地とは話者が０～１０歳までで、一番長くいた県、州、又はそれに類する地名を指すこととし、それを 2 byte コードで記述する。

＜例＞ 発呼者: 北野・やよい_ 女性_ 20 _ 大阪_ アナウンサー

5.3.18　被呼者

電話を受けた方の話者（被呼者の話者ＩＤ＜Ｂ＞にあたる人物）の情報を記述する。
発呼者と同一の『話者の名前_ 性別_ 年齢_ 出身地_ 職業』のフォーマットで記述する。

＜例＞ 被呼者: 村上・仁一_ 男性_ 19 _ 北海道_ 会社員

5.3.19　通訳者 1

発呼者と被呼者が別の言語を話した場合、その中間に立って通訳した人物の情報を記述する。
通訳者の話者ＩＤ＜Ｃ＞にあたる人物の情報である。
発呼者と同一の『話者の名前_ 性別_ 年齢_ 出身地_ 職業』のフォーマットで記述する。

＜例＞ 通訳者 1 : キャンベル・ニック_ 男性_ 35 _ スコットランド_ 通訳

5.3.20　通訳者２

通訳を２名で行なった場合、もう一人の通訳者についての情報を記述する。通訳者の話者ＩＤ＜Ｄ＞にあたる人物の情報である。

発呼者と同一の『話者の名前_ 性別_ 年齢_ 出身地_ 職業』のフォーマットで記述する。

＜例＞ 通訳者２: 田代・敏久_ 男性_ 40 _ 北海道_ 通訳

5.3.21　翻訳者

発呼者と被呼者が同一の言語で話した時、その会話を後で翻訳した場合に、その翻訳を行なった人物の情報を記述する。

発呼者と同一の『話者の名前_ 性別_ 年齢_ 出身地_ 職業』のフォーマットで記述する。

＜例＞ 翻訳者: 浦谷・則好_ 男性_ 50 _ 沖縄_ 会社員

5.4 追記事項

ここで作成される会話属性情報は各種プログラムにかけられる可能性があるので、いくつかの注意点を挙げておく。

1. このファイルでの漢字コードはＥＵＣコードを用いる。
2. ヘッダーに使用されている「:」は 1 byte コードに統一する。
3. 1 byte、2 byte を問わず、スペースは使用しない。
4. タブは使用しない。
5. 名前などで、カタカナ表記又は英字表記を行なう場合、同一人物の名前の表記は複数会話に渡って統一をはかる。
6. 「通訳者:」「翻訳者:」「コメント:」など、情報が存在しない又は記入不可能な場合は記入すべきフィールドに 1 byte コードで「NIL」と記述する。

会話属性情報（会話ＩＤ.ENV）の作成例

```
会話ＩＤ:SXS00001
収録日時:1969 年 10 月 04 日
収録社名: ＡＴＲ音声翻訳通信研究所
総会話時間:78 分 21 秒
有効会話時間:00 分 21 秒
ノイズレベル:45dBA ＜無響音室＞
マイクロホン: ＭＵ－２Ｃ＜ｓａｎｋｅｎ＞
ＤＡＴの機種: ＸＤ－Ｚ５０５＜Ｖｉｃｔｏｒ＞
ミキサーの機種: Ａ－８１２ＥＸ＜ＯＮＫＹＯ＞
サンプリング周波数:48kHz
発話形態: 自由発話
ドメイン: 会議の日程調整
トピック:NIL
言語パターン: 日本語－日本語
発声方法: 文発声
対面・非対面: 非対面
発呼者: 北野・やよい_ 女性_ 20 _ 大阪_ アナウンサー
被呼者: 村上・仁一_ 男性_ 19 _ 北海道_ 会社員
通訳者 1 :NIL
通訳者 2 :NIL
翻訳者:NIL
コメント:NIL
```

6．おわりに

音声言語データベースはまだ構築を開始したばかりである。ここで紹介した仕様も後に変更になる部分があるかもしれない。しかし、データベースは利用されてこそ価値を発揮するものであることに違いはない。この報告書はそのために少しでも早く利用者に音声言語データベースに関する情報を提供することを目的に書かれたものである。筆者等は少しでも早く多くのデータを提供し、このデータベースを用いた成果が少しでも早く出てくることを期待している。

本格的なデータベースの構築は少人数でできるものではない。この音声言語データベースの構築にも多くの人々が関わっている。特に、「データベース作業部会」は中心となってこの作業を進めている。メンバーは浦谷，竹沢ほか松永昭一主任研究員，村上仁一研究員，樋口宜男室長（第２研究室），Nick Campbell主幹研究員，古瀬蔵主任研究員，側嶋康博研究員の８名である。（特に、データ収集に関しては古瀬主任研究員，側嶋研究員が中心に実施している。）また、方針の決定に当たっては山崎泰弘社長，森元逞室長（第４研究室），飯田仁室長（第３研究室），匂坂芳典室長（第１研究室）には多大なご協力をいただいている。田代敏久研究員には日本語情報の付与に関して，赤峯享研究員，傳康晴研究員には英語情報付与に関して中心となって基準の作成に当たっていただいた。日本ＩＲの衛藤純司さん，坂口明子さんやＴＩＳの田中ゆかりさん，山田久子さんなどメーカーの方々にも色々とご協力いただいた。この報告書はこのように様々な人々の協同作業の結果である。音声言語データベース作成に関わった全ての人々に感謝する。

参考文献

[1]江原暉将，小倉健太郎，篠崎直子，森元 逞，樽松 明：電話またはキーボードを介した対話に基づく対話データベースＡＤＤの構築，情報処理学会論文誌，Vol.33，No.4,pp.448-456(1992)

[2]武田一哉，匂坂芳典，片桐 滋，桑原尚夫：研究用日本語音声データ・ベースの構築，日本音響学会誌，Vol.44,No.10,pp.747-754(1988)

[3]匂坂芳典，浦谷則好：ＡＴＲ音声・言語データベース，日本音響学会誌，Vol.48,No.12,pp.878-882(1992)

[4]浦谷則好，竹沢寿幸，田代敏久，森元 逞，匂坂芳典：ＡＴＲの新音声言語データベース，情報処理学会第４８回全国大会 3-3Q-4(1994)

[5]浦谷則好，竹沢寿幸，田代敏久，衛藤純司：音声言語データベースのための日本語形態素情報と表記の体系，ATR Technical Report TR-IT-0003(1994)

[6]浦谷則好，田代敏久，山田久子，松本 香：音声言語データベースにおける日本語形態素解析マニュアル，ATR Technical Report TR-IT-0009(1994)