TR-I-0371

ＡＴＲ対話データベース検索内容調査

A Sarvey of Data Retrieval from the ATR Dialogue Database

浦谷則好　　井上健吾*
Noriyoshi URATANI　Kengo INOUE

1993.3

概要

数々の言語現象を調査研究する目的でＡＴＲ対話データベース（ＡＤＤ）に対して種々の検索が行なわれている。データベースの見直しの資料とするために検索内容の傾向について調査をした。調査項目は、検索の種類、利用したデータとその件数である。期間は１９９０年４月から１９９２年１２月末を対象とした。

調査結果からは次のことが明らかになった。検索対象になったデータは国際会議のタスク（領域）を指定しているものが圧倒的に多く、旅行のタスクのみを指定した検索条件は少数である。キーボード会話、電話会話等のメディアの選択は双方ともほぼ同じ頻度で利用されている。利用データとしては、調査対象とした検索の殆どが何らかの日本語形態素情報を利用している。日本語形態素情報に合わせて対応英語を求めている検索も多く、対応単位としては文対応が大勢を占める。利用データとして英語形態素情報が付与されているデータがＡＤＤに反映している量が少なく、また低品質であったため、英語表現を基準とした検索は殆ど見受けられなかった。

ATR 自動翻訳電話研究所
ATR Interpreting Telephony Research Laboratories

*株式会社 東洋情報システム
TOYO Information Systems

目次

１．はじめに

ＡＴＲ対話データベース（以下ＡＤＤに略する）に対して各種検索が行なわれてきた。データベースの見直しの資料として検索内容の傾向について調査した。

２．調査内容

検索内容の調査項目は以下の通りである。

（１）検索の種類
メディア、タスク（領域）、会話ＩＤ指定等の検索を行なう対象のデータが、どの様な種別であるかを調査した。

（２）利用したデータ
日本語形態素や係り受けデータ、日英対応データ等のＡＤＤに登録される上で付与されたデータのどの部分を必要としている検索か。また組合せているかを調査した。ただし、日英対応データを利用している場合、どのレベルが必要であるかまで調査した（文対応か文節対応それともランダム対応が必要である等）。また、単に文字化データのみで、形態素データさえ不要な検索の場合は「原データのみ」とする。

（３）頻度調査
調査項目の「検索の種類」と「利用したデータ」の組合せて検索件数の頻度を調査した。
ただし、類似した検索内容で、後の検索内容が前の検索内容の修正と思われる場合には、前の検索内容を件数には含めない。

*1 メディア：
電話、キーボード、モデル

*2 タスク：
国際会議、旅行

３．調査の対象とする期間

１９９０年４月～１９９２年１２月末

４．調査対象の検索数

総数　：　７８件　【期間中に作業を行なった検索数】
修正検索：　３件
調査対象：　７５件

５．調査結果

（１）検索の種類

検索の領域やメディア等の指定状況

(01)　１９件　領　　域:国際会議
　　メディア:電話,キーボード

(02)　１４件　領　　域:国際会議,旅行
　　メディア:電話,キーボード

(03)　８件　領　　域:国際会議
　　メディア:キーボード

(04)　５件　領　　域:国際会議
　　メディア:電話

(05)　３件　領　　域:国際会議,旅行
　　メディア:電話

(06)　３件　領　　域:国際会議,旅行
　　メディア:キーボード

(07)　２件　領　　域:国際会議,旅行
　　メディア:電話,キーボード
　　内　　容:問合せ
　　言語パターン:日日

(08)　２件　全会話（利用できるもの全て）

(09)　２件　会話ＩＤ:81～129

(10)　１件　領　　域:旅行
　　メディア:電話,キーボード

(11)　１件　領　　域:国際会議
　　メディア:電話
　　内　　容:問合せ
　　言語パターン:日日

(12)　１件　領　　域:国際会議
　　メディア:電話
　　内　　容:問合せ
　　言語パターン:日英

(13)　１件　領　　域:国際会議
　　メディア:電話
　　内　　容:問合せ
　　言語パターン:英日,日英

(14)　１件　領　　域:国際会議
　　メディア:電話
　　内　　容:問合せ

(15)　１件　領　　域:国際会議
　　メディア:電話
　　言語パターン:日英

(16)　１件　領　　域:国際会議
　　メディア:電話
　　言語パターン:英日,日英,英日（同時通訳）

(17)　1件 領　域:国際会議
メディア:電話
会話ＩＤ:1,5,6,15～19

(18)　1件 領　域:国際会議
メディア:電話
会話ＩＤ:1～26,28,30,32,34～40,58,59,60,61

(19)　1件 領　域:国際会議
メディア:電話,キーボード
内　容:問合せ
言語パターン:英日,日英

(20)　1件 領　域:国際会議
メディア:電話,キーボード
言語パターン:日日

(21)　1件 領　域:国際会議
メディア:電話,キーボード
言語パターン:日英

(22)　1件 領　域:国際会議,旅行
メディア:電話
内　容:問合せ

(23)　1件 領　域:国際会議,旅行
メディア:電話,キーボード
言語パターン:英日,日英

(24)　1件 領　域:国際会議,旅行
メディア:キーボード
言語パターン:英日,日英,日日

(25)　1件 会話ＩＤ:81～129,3045～3142

(26)　1件 会話ＩＤ:81

（２）利用したデータ

検索で利用した情報の種類
実際の検索例として付録１に検索作業ドキュメントを添付する（１～１２の番号に対応）。

(01)　３２件 形態素情報

(02)　１９件 形態素情報
日英対応情報:文対応

(03)　８件 形態素情報
係り受け情報

(04)　４件 原データのみ

(05)　３件 形態素情報
日英対応情報:発話対応

(06)　２件 形態素情報
日英対応情報:単語対応
係り受け情報

(07)　２件 形態素情報
日英対応情報:単語対応,文対応

(08)　１件 形態素情報
日英対応情報:文節対応
係り受け情報

(09)　１件 形態素情報
日英対応情報:文節対応

(10)　１件 形態素情報
日英対応情報:単語対応

(11)　１件 形態素情報
係り受け情報
日英対応情報:格対応

(12)　１件 各種会話情報

（３）「検索の種類」と「利用したデータ」を組み合せた件数

検索の種類と利用した情報を組み合わせで件数を算出

(01)　１０件　種　類|領　域:国際会議
メディア:電話,キーボード
データ|形態素情報

(02)　６件　種　類|領　域:国際会議,旅行
メディア:電話,キーボード
データ|形態素情報

(03)　５件　種　類|領　域:国際会議
メディア:電話,キーボード
データ|形態素情報
係り受け情報

(04)　４件　種　類|領　域:国際会議
メディア:キーボード
データ|形態素情報
日英対応情報:文対応

(05)　３件　種　類|領　域:国際会議
メディア:電話
データ|形態素情報
日英対応情報:文対応

(06)　２件　種　類|領　域:国際会議
メディア:電話,キーボード
データ|形態素情報
日英対応情報:文対応

(07)　２件　種　類|領　域:国際会議
メディア:キーボード
データ|形態素情報

(08)　２件　種　類|領　域:国際会議,旅行
メディア:電話
データ|形態素情報

(09)　２件　種　類|領　域:国際会議,旅行
メディア:電話,キーボード
データ|形態素情報
日英対応情報:文対応

(10)　２件　種　類|領　域:国際会議,旅行
メディア:電話,キーボード
データ|形態素情報
係り受け情報

(11)　２件　種　類|領　域:国際会議,旅行
メディア:電話,キーボード
内　容:問合せ
言語パターン:日日
データ|原データのみ

(12)　２件　種　類|領　域:国際会議,旅行
メディア:キーボード
データ|形態素情報

(13)　２件　種　類|全会話（利用できるもの全て）
データ|形態素情報

(14) 1件 種 類|領 域:旅行
メディア:電話,キーボード
データ|形態素情報

(15) 1件 種 類|領 域:国際会議
メディア:電話
データ|原データのみ

(16) 1件 種 類|領 域:国際会議
メディア:電話
データ|形態素情報

(17) 1件 種 類|領 域:国際会議
メディア:電話
内 容:問合せ
データ|形態素情報
日英対応情報:文対応

(18) 1件 種 類|領 域:国際会議
メディア:電話
内 容:問合せ
言語パターン:日日
データ|形態素情報

(19) 1件 種 類|領 域:国際会議
メディア:電話
内 容:問合せ
言語パターン:日英
データ|形態素情報

(20) 1件 種 類|領 域:国際会議
メディア:電話
内 容:問合せ
言語パターン:英日,日英
データ|形態素情報
日英対応情報:文対応

(21) 1件 種 類|領 域:国際会議
メディア:電話
言語パターン:日英
データ|形態素情報

(22) 1件 種 類|領 域:国際会議
メディア:電話
言語パターン:英日,日英,英日(同時通訳)
データ|形態素情報
日英対応情報:発話対応

(23) 1件 種 類|領 域:国際会議
メディア:電話
会話ID:1,5,6,15～19
データ|形態素情報
日英対応情報:文対応

(24) 1件 種 類|領 域:国際会議
メディア:電話
会話ID:1～26,28,30,32,34～40,58,59,60,61
データ|形態素情報

(25) 1件 種 類|領 域:国際会議
メディア:電話,キーボード
データ|形態素情報
日英対応情報:発話対応

(26) 1件 種 類|領 域:国際会議
メディア:電話,キーボード

データ|形態素情報
日英対応情報:単語対応,文対応

(27) 1件 種 類|領 域:国際会議
メディア:電話,キーボード
内 容:問合せ
言語パターン:英日,日英
データ|形態素情報
日英対応情報:文対応

(28) 1件 種 類|領 域:国際会議
メディア:電話,キーボード
言語パターン:日日
データ|原データのみ

(29) 1件 種 類|領 域:国際会議
メディア:電話,キーボード
言語パターン:日英
データ|形態素情報
日英対応情報:文対応

(30) 1件 種 類|領 域:国際会議
メディア:キーボード
データ|形態素情報
日英対応情報:文節対応
係り受け情報

(31) 1件 種 類|領 域:国際会議
メディア:キーボード
データ|形態素情報
日英対応情報:発話対応

(32) 1件 種 類|領 域:国際会議,旅行
メディア:電話
データ|形態素情報
日英対応情報:文対応

(33) 1件 種 類|領 域:国際会議,旅行
メディア:電話
内 容:問合せ
データ|形態素情報
日英対応情報:文対応

(34) 1件 種 類|領 域:国際会議,旅行
メディア:電話,キーボード
データ|形態素情報
日英対応情報:単語対応
係り受け情報

(35) 1件 種 類|領 域:国際会議,旅行
メディア:電話,キーボード
データ|形態素情報
日英対応情報:単語対応,文対応

(36) 1件 種 類|領 域:国際会議,旅行
メディア:電話,キーボード
データ|形態素情報
日英対応情報:単語対応

(37) 1件 種 類|領 域:国際会議,旅行
メディア:電話,キーボード
データ|各種会話情報

(38) 1件 種 類|領 域:国際会議,旅行
メディア:電話,キーボード
言語パターン:英日,日英

データ|形態素情報

(39) 1件 種　類|領　　域:国際会議,旅行
メディア:キーボード
データ|形態素情報
係り受け情報

(40) 1件 種　類|領　　域:国際会議,旅行
メディア:キーボード
言語パターン:英日,日英,日日
データ|形態素情報
日英対応情報:文対応

(41) 1件 種　類|会話ＩＤ:81
データ|形態素情報
日英対応情報:単語対応
係り受け情報

(42) 1件 種　類|会話ＩＤ:81～129
データ|形態素情報
日英対応情報:文節対応

(43) 1件 種　類|会話ＩＤ:81～129
データ|形態素情報

(44) 1件 種　類|会話ＩＤ:81～129,3045～3142
データ|形態素情報
係り受け情報
日英対応情報:格対応

付録１ 調査項目「利用したデータ」の検索例

(01)　　３２件 形態素情報

```
**********************************************************************************
                          検索依頼ドキュメント
**********************************************************************************
  １．依 頼 者： 保坂研究員

  ２．依頼日時： 90/06/07

  ３．回答日時： 90/07/19

  ４．回 答 者： 高橋.井上.庄山 (TIS)

  ５．内　　容： 文節における助詞及び助動詞の連接状況。

  ６．対象会話： キーボード/電話(全ての会話)

  ７．出力形式： 出現頻度 品詞/品詞/...|読み/読み/...

  ８．検索手順： 下欄に別記する。

  ９．ファイル： 使用ファイルの一覧　($SEARCH = /data6/LDB/search)
        検索結果：
          /data6/LDB/search/ans/FURUSE/FURUSE4/HOSAKA2.key
                                              /HOSAKA2.tel

        個別プログラム：　($SEARCH/pgmにも存在する。)
          $SEARCH/work/HOSAKA/HOSAKA2/get_jwrd
                                     /jyodousi-jyosi_convert
                                     /jyodousi-jyosi_convert.sub1
                                     /jyodousi-jyosi_convert.sub2

１０．内容補足：・連接状態別に出現頻度を出力する。

１１．備　　考：・助詞、助動詞の連接は記号、自立語、その他を省いた
                 品詞と助詞、助動詞の連接とする。

１２．問 題 点：・ＬＤＢＳＨでは文節をテキスト化出来るが、その中に含まれる
                 単語数は特定出来ないために、品詞を同時に出力する事が
                 出来ない。
                ・単語テーブルに、データの抜けている部分があり出力結果に
                 不備がある。

１３．依頼ＩＤ：HOSAKA2

**********************************************************************************
```

検索手順

１．文節を作るために全ての単語を検索。（ＳＱＬ）

ex 3001|　　100|もしもし|11
3001|　　100|、|00
3001|　　200|通訳電話国際会議事務局|30
3001|　　200|です|12

２．単語を会話ＩＤと文節ＩＤをＫＥＹとして文節にする。

ex 3001-100|もしもし:11/、:00
3001-200|通訳電話国際会議事務局:30/です:12/か:16/？:00

３．文節データのファイルから助詞、助動詞を含む文節を抽出。

ex 3001-200|通訳電話国際会議事務局:30/です:12/か:16/？:00

４．品詞コードを品詞名にし、文節を助詞、助動詞の連接のパターンに、
加工する。

ex 助動詞/終助詞|です/か

５．出現頻度を算出する。

ex 200 助動詞/終助詞|です/か

以上

****
検索依頼ドキュメント
****

１．依 頼 者： 工藤 研究員

２．依頼日時： 90/11/20

３．回答日時： 90/11/26

４．作業日数： ２ 日間

５．回 答 者： 井上・庄山 （株）東洋情報システム

６．内　　容： 代名詞、接頭語「何」を含む疑問文を出力

７．対象会話： 領　　域: 国際会議／旅行
メディア: キーボード

８．出力形式： 会話ＩＤ-文ＩＤ|日本語文

９．出力サンプル： 下欄に別記する。

１０．検索手順： 下欄に別記する。

１１．ファイル： 使用ファイルの一覧 ($SEARCH = /data/LDB/search)

検索結果:
$SEARCH/ans/KUDOU/KUDOU3/gimon.daimei
/gimon.settou

個別プログラム:
$SEARCH/work/KUDOU/KUDOU3/KUDOU3.bat

データソース:
/data6/LDB/search/lib/TABLES/j_word.idx
/j_phrase.idx

データ加工用ツール:
$SEARCH/tool/select-from.csh
/jwrd_to_sent

１２．内容補足： 検索結果は代名詞の場合と接頭語の場合とでファイルを分けています

１３．備　　考：

１４．依頼ＩＤ：KUDOU3

****
#出力サンプル
・gimon.daimei
会話ＩＤ-文ＩＤ|日本語文

3001-1100|もし、何か分からない点が、ございましたら、こちらの方へ、いつでもお聞き下さい。
3003-3000|その費用には何が含まれていますか？
3003-3500|それと発表者が、スライドを使いたいと言っていますが、何か機械等ありますか？
3005-400|はい、何でしょうか？
3006-300|ちょっとお尋ねしますが、今回の会議の目的は何ですか？
3007-2600|会議は何語で運営されるのですか。
3008-3500|他に何か？
3009-2600|他に何か御質問は？
3010-200|もしもし、御用件は何でしょうか？
3010-1500|他に何か？

****
検索依頼ドキュメント
****

１．依 頼 者： 浦谷 主幹研究員

２．依頼日時： 91/10/18

３．回答日時： 91/10/30

４．作業日数： ６日間

５．回 答 者： 井上 健吾 （株）東洋情報システム

６．内　　容： １文節内における名詞連続の出現頻度を出力

７．対象会話： 領　　域: 国際会議
メディア: 電話/キーボード

８．出力形式： 出現頻度|後名詞|前名詞

９．出力サンプル： 下欄に別記する。

１０．検索手順： 検索別バッチプログラム参照（別冊添付資料）

１１．ファイル： 使用ファイルの一覧
($SEARCH = /data/LDB/search, $TBMT = /data6/LDB/TBMT/KB)

検索結果:
$SEARCH/ans/URATANI/URATANI3/2_nouns.ans

バッチプログラム:
$SEARCH/work/URATANI3.bat

データソース:
$SEARCH/work/TAKEZAWA/TAKEZAWA3/jword.mast

データ加工用ツール:
$SEARCH/tool/

$TBMT/tool/

１２．内容補足：なし

１３．備　　考：なし

１４．依頼ＩＤ：URATANI3

****

＃出力サンプル

ファイル名：
atr-dp:/data6/LDB/search/ans/URATANI/URATANI3/2_nouns.ans

出力形式：
出現頻度|後名詞|前名詞

検索結果：
343|局|事務
197|料|登録
94|番号|電話
89|場|会議
81|登録|参加
63|用紙|登録
54|会|分科
52|用紙|申込み
49|委員|企画
46|会|委員
40|料|参加
40|番号|郵便
37|申込み|参加
31|会議|国際
30|先|連絡
29|講演|基調
29|後|終了
28|通訳|同時
28|終了|会議
28|委員|組織
26|長|委員
24|市|見本
23|参加|会議
23|講演|招待
22|参加|一般
21|録|議事
21|知能|人工
20|ツアー|観光
20|カード|登録
18|翻訳|機械
18|振込み|銀行
17|方法|支払
17|番号|ファックス
17|集|論文
17|室|研究
17|関係|大学
16|集|要旨
16|事項|必要
15|料|手数
15|当日|会議
15|電話|通訳
15|講演|記念
15|応答|質疑
14|式|開会
13|日|最終
13|日|開催
13|受付|登録
13|公立|国
12|文|挨拶
12|部会|専門

以上

****
検索依頼ドキュメント
****

１．依 頼 者： 浦谷 主幹研究員

２．依頼日時： 92/03/04

３．回答日時： 92/03/30

４．作業日数： ４日間

５．回 答 者： 井上 健吾 （株）東洋情報システム

６．内　　容： サ変名詞と「有り」が、名詞と連続するパターンを抽出し頻度を出力
普通名詞＋（接尾語）＋サ変名詞
普通名詞＋（接尾語）＋「有る」（標準形）

７．対象会話： 領　　域: 国際会議
メディア: 電話／キーボード

８．出力形式： 名詞文字列/（サ変名詞|「有り」）|頻度

９．出力サンプル： 下欄に別記する。

１０．検索手順： 検索別バッチプログラム参照 （別冊添付資料）

１１．ファイル： 使用ファイルの一覧
($SEARCH = /data/LDB/search, $TBMT = /data6/LDB/TBMT/KB)

検索結果:
$SEARCH/ans/URATANI/URATANI9/noun-sahen.ans

バッチプログラム:
$SEARCH/work/URATANI/URATANI9/URATANI9.bat

データソース:
ＡＴＲ対話データベース

データ加工用ツール:
$SEARCH/tool/kifilter
c2p.gawk
$TBMT/tool/REFORM

１２．内容補足： 言い淀み言い直しを削除した後に検索を行なう。

１３．備　　考： 特に無し

１４．依頼ＩＤ： URATANI9

****

＃出力サンプル

申訳/有る|61
問題/有る|16
質問/有る|8
必要/有る|7
仕方/有る|5
方/用意|3
色色/有る|3
所/予定|3
興味/有る|3
関係/有る|3
間違/有る|3
凡て/手配|2
多数/列挙|2
色色/検討|2
色色/アレンジ|2
十分/有る|2
事/有る|2
再度/確認|2
差支え/有る|2
今/予定|2
今回/参加|2
後日/連絡|2
ホテル/有る|2
連絡/有る|1
両方/入手|1
両方/共/購入|1
両方/とも/実施|1
来年/開催|1
来週/発送|1
来月/開催|1
余裕/有る|1
予約/発注|1
凡て/収録|1
翻訳/者/とも/リスト・アップ|1
方/復誦|1
方/登録|1
方/担当|1
方/準備|1
方/検討|1
分野/カバー|1
物/郵送|1
物/有る|1
風/諒解|1
部屋/予約|1
必要/事項/記入|1
彼方此方/観光|1
半/終了|1
入金/確認|1
日/講演|1
二十五日/到着|1

以上

(02)　　１９件 形態素情報
　　日英対応情報:文対応

****
検索依頼ドキュメント
****

１．依 頼 者： 友清 研究員

２．依頼日時： 91/02/20

３．回答日時： 91/04/03

４．作業日数： １３日間

５．回 答 者： 井上 健吾 （株）東洋情報システム

６．内　　容： 検索パタンを含む日本語文とその対訳英語文を出力

検索パタン:
「下さる（くださる）」 連用形（下さい、ください）、
命令形（下さい、ください）
「くれる」 連用形・命令形（くれ）、終止形（くれる）
「もらう」 連用形（もらえ）、終止形（もらう）
「いただける」 連用形（いただけ）、終止形（いただける）
「ほしい」 連体形（ほしい）
「たい」 連体形（たい）
「いい」 連体形（いい）

７．対象会話： 領　　域: 国際会議
メディア: 電話
内　　容: 問合せ

８．出力形式： 会話ＩＤ|文対応ＩＤ|日本語文ＩＤ@日本語文(_日本語文ＩＤ@¥
日本語文)|英語文ＩＤ@英語文(_英語文ＩＤ@英語文)

９．出力サンプル： 下欄に別記する。

１０．検索手順： 検索別バッチプログラム参照 （別冊添付資料）

１１．ファイル： 使用ファイルの一覧
($SEARCH = /data6/LDB/search, $TBMT = /data6/LDB/TBMT/KB)

検索結果:
$SEARCH/ans/TOMOKIYO/TOMOKIYO7/hosii.ans
/ii.ans
/itadake.ans
/itadakeru.ans
/kudasai_mei.ans
/kudasai_ren.ans
/kure.ans
/kureru.ans
/morae.ans
/morau.ans
/tai.ans

バッチプログラム:
$SEARCH/work/TOMOKIYO/TOMOKIYO7/TOMOKIYO7.bat

データソース:
文節情報ファイル
$TBMT/corpus/SENTENCE.E

データ加工用ツール:
$SEARCH/tool/MKJSENT

/kifilter2

$TBMT/tool/JOIN
/CKSORT
/RECONSTRUCT

１２．内容補足：文対応単位で日本語文と英語文を１レコードとしているので
日英文とも複数文存在するものが有る。

１３．備　　考：なし

１４．依頼ＩＤ：TOMOKIYO7

****
#出力サンプル

ファイル:
atr-dp:/data6/LDB/search/ans/TOMOKIYO/TOMOKIYO7/hosii.ans

出力形式:
会話ＩＤ|文対応ＩＤ|日本語文ＩＤ@日本語文(_日本語文ＩＤ@日本語文)|¥
英語文ＩＤ@英語文(_英語文ＩＤ@英語文)

検索結果:
27|500|600@そちらから会場案内の地図と市バスの時刻表頂いてるんですけれども、これちょっとわかりにくいので説明してほしいんですけれども。|1000@Well , you enclosed the map of Kyoto and also the timetable for bus ,_1200@but these aren't very easy to read , so I'd like you to explain a little bit for me , if you don't mind .
41|500|1700@はい、ということは、問い合わせの方の方が、コールフォーペーパーというのを、お持ちなのかということを、ちょっとお伺いしてほしいんですけれども。|2700@Have you a ** inaudible ** call-for paper ?
177|200|200@アカンパニー・パーソンのことについて教えてほしいんです。|300@I want to know about ' accompany person ' .

以上

****

検索依頼ドキュメント

****

１．依 頼 者： 友清研究員

２．依頼日時： 91/06/14

３．回答日時： 90/06/14

４．作業日数： １日間

５．回 答 者： 井上 健吾 （株）東洋情報システム

６．内　　容： 検索パタンを含む日本語文とその対訳英文を出力

検索パタン：
「ほど」、「やっぱり」、「やはり」、「年上」、「年下」

７．対象会話： 領　　域: 国際会議
メディア: 電話
内　　容: 問合せ

８．出力形式： 会話ＩＤ|文対応ＩＤ|日本語文ＩＤ@日本語文(_日本語文¥
ＩＤ@日本語文)|英語文ＩＤ@英語文(_英語文ＩＤ@英語文)

９．出力サンプル： 下欄に別記する。

１０．検索手順： 検索別バッチプログラム参照 （別冊添付資料）

１１．ファイル： 使用ファイルの一覧
($SEARCH = /data/LDB/search, $TBMT = /data6/LDB/TBMT/KB)

検索結果:
$SEARCH/ans/TOMOKIYO/TOMOKIYO8/hodo.ans
/tosisita.ans
/tosiue.ans
/yahari.ans
/yappari.ans

バッチプログラム:
$SEARCH/work/TOMOKIYO/TOMOKIYO8/TOMOKIYO8.bat

データソース:
$TBMT/corpus/SENTENCE.JE

データ加工用ツール:
$SEARCH/tool/PATTERN_MATCH
/pickup_data
/kifilter2

$TBMT/tool/

１２．内容補足： 検索パターン、「年上」、「年下」は検索出来ませんでした。

１３．備　　考：なし

１４．依頼ＩＤ：TOMOKIYO8.bat

****

＃出力サンプル

ファイル：
atr-dp:/data6/LDB/search/ans/TOMOKIYO/TOMOKIYO8/hodo.ans

出力形式：
会話ＩＤ|文対応ＩＤ|日本語文ＩＤ@日本語文(_日本語文ＩＤ@日本語文)|¥
英語文ＩＤ@英語文(_英語文ＩＤ@英語文)

検索結果：
27|1100|2000@それから京都駅から会場までタクシーでずっと来ていただいても、_2200@１，０００円ぐらいですので、それほど高くないと思いますが。|2000@You could also take a taxi ._2100@It's not too expensive , about 1,000 < a thousand > yen , I think .
37|2500|10200@いえ、タクシーで２分ほどになっております。|12800@Not so far , about 2 < two > minutes by taxi .
38|1800|9300@電車で１時間ほどです。|7400@About 1 < one > hour by train .
38|6600|35100@そうですね、お酒とか料理のお金が、１晩で大体３０，０００円ぐらいですね。_35300@あと芸者さんとか頼むお金が、７０，０００円から１００，０００円ほど、ということです。|40100@It will take 30,000 < thirty thousand > yen for drinking ,_40300@and 70 < seventy > to 100 < one hundred > for Geisha .
38|7500|40900@それで１時間ほどたちまして、芸者さんが来られて、_41100@で、いろいろと接待をしてくれる、ということです。|45100@And 1 < one > hour later , the Geisha-san will come ,_45300@and talk to you .
38|7800|41900@私も、それほど上手ではないんですけれども、_42100@一応、片言ならしゃべれますんで、_42300@お役に立てると思いますけれど。|46000@Well , I'm not so good at English , but I can be helpful of you .
42|1100|3000@電車で１時間ほど、ですが。|5200@It takes about 1 < one > hour by train .

以上

****

検索依頼ドキュメント

****

１．依 頼 者： 友清 研究員

２．依頼日時： 92/03/04

３．回答日時： 92/03/19

４．作業日数： ９日間

５．回 答 者： 井上 健吾 （株）東洋情報システム

６．内　　容： 名詞（名詞連続を含む）の直後が助詞でないものをすべてと
その対訳を出力。

７．対象会話： 領　　域: 国際会議
メディア: 電話
内　　容: 問合せ
言語パターン: 日英

８．出力形式： 出現単語（名詞類）|品詞名
<TAB>日本語文
<TAB>英語文

９．出力サンプル： 下欄に別記する。

１０．検索手順： 検索別バッチプログラム参照 （別冊添付資料）

１１．ファイル： 使用ファイルの一覧
($SEARCH = /data/LDB/search, $TBMT = /data6/LDB/TBMT/KB)

検索結果:
$SEARCH/ans/TOMOKIYO/TOMOKIYO/no-particle_noun

バッチプログラム:
$SEARCH/work/TOMOKIYO/TOMOKIYO/TOMOKIYO12.bat

データソース:
ＡＴＲ対話データベースを使用

データ加工用ツール:
$SEARCH/tool/kifilter
/c2p.gawk

$TBMT/tool/REFORM
/RECONSTRUCT
/JOIN2

１２．内容補足： 検索結果は品詞名でソートしてあります。

１３．備　　考： 特に無し

１４．依頼ＩＤ： TOMOKIYO12

****

＃出力サンプル

ファイル名： no-particle_noun

ＰＲ|サ変名詞
是非、私共もこの会議を世界にＰＲしたいと思っておるんですが。
We are also making people at large know more about this conference, too.

あいさつ|サ変名詞
うーんでは、会議で冒頭であいさつする程度ですね。
Oh, I guess then, they make sort of an opening or welcoming speech at the outset of the conference, I guess.

はっきり|サ変名詞
いくつかちょっとはっきりしない点があるのでもう一回お聞きしたいんですが。
まず、カメラについてですけれども、先程は確かフラッシュが、スピーカーの人たちのじゃまになるので使ってはいけないというふうにおっしゃったと思うんですが。
I'm not quite clear on the point regarding the cameras.
you specifically mentioned flash earlier and I know the flash does disturb the speakers.

らん|サ変名詞
ですから、その際には、ダイヤの方を正確にごらんいただいて、まちがった時間表をごらんにならないように、御注意下さい。
So make sure to read the right and correct bus timetable.

アレンジ|サ変名詞
そしたらわざわざインフォメーション送って頂かなくても結構です。
実は私たちは３日以上続くツアーを探していましたので、それでしたらこちらの方の旅行社を使って自分でアレンジするか、あるいは日本に着いてからアレンジするかにしたいと思います。
Well, then I don't think you need to send me that information because we are looking for something that's longer than three days.
So I think I'll try and arrange something either through my travel agent here or after I get to Japan.

アレンジ|サ変名詞
そしたらわざわざインフォメーション送って頂かなくても結構です。
実は私たちは３日以上続くツアーを探していましたので、それでしたらこちらの方の旅行社を使って自分でアレンジするか、あるいは日本に着いてからアレンジするかにしたいと思います。
Well, then I don't think you need to send me that information because we are looking for something that's longer than three days.
So I think I'll try and arrange something either through my travel agent here or after I get to Japan.

アレンジ|サ変名詞
ちなみにですねホテルニューオータニ大阪の横には大阪城ホールという見本市会場があるんですけども、今回一応国際コンピューター会議は一応見本市を併設してませんので、こちらの方では要するに一般的な見本市としての見本市が併催行事されてる国際会議ではないので、今んところ御社を含めてですね、２・３社しか出展あるいは書籍展示等しか一応こちらの方にはでてませんのでとりあえずホテル内で私共のアレンジする必要があるかと思います。
for this conference we will use the Hotel New Otani Osaka, and we have an exhibition hall called the Osaka Castle Hall near the hotel.
However, this International Conference on Computer Science does not have any trade fairs.
Therefore, there are only 2, 3 companies who would like to exhibit their products or the books, including you.
So we would like to use the hotel accommodations so that we can have the exhibitions or exhibits.

以上

****
検索依頼ドキュメント
****

１．依 頼 者： 隅田 研究顕

２．依頼日時： 92/06/02

３．回答日時： 作業保留中

４．作業日数：

５．回 答 者： 井上 健吾 （株）東洋情報システム

６．内　　容： ”ｓｅｎｄ Ｗ１ Ｗ２ … Ｗｎ ｔｏ”を含む英文と対訳日本語文を出力。
Ｗ ＝ 単語、 ”ｎ”は１０以内

７．対象会話： 領　　域: 国際会議
メディア: キーボード

８．出力形式： 会話ＩＤ| 文対応ＩＤ
<TAB>文ＩＤ@英語文
<TAB>文ＩＤ@日本語文

９．出力サンプル： 下欄に別記する。

１０．検索手順： 検索別バッチプログラム参照 （別冊添付資料）

１１．ファイル： 使用ファイルの一覧
($SEARCH = /data/LDB/search, $TBMT = /data6/LDB/TBMT/KB)

サンプル:
$SEARCH/sample/SUMITA/SUMITA6/sample.ans

バッチプログラム:
$SEARCH/work/SUMITA/SUMITA6/SUMITA6.bat

データソース:
テキスト化ＡＤＤを使用

データ加工用ツール:
$SEARCH/tool/kifilter
/mkengsent.gawk
$TBMT/tool/REFORM
/RECONSTRUCT
/JOIN2

１２．内容補足： 特に無し

１３．備　　考：

１４．依頼ＩＤ：SUMITA6

****

#出力サンプル

ファイル名： sample.ans

3001|600
600@Okay, then I will [send] them [to] you.
700@Would you please give me your name and address?
600@分りました。
700@それでは、こちらからお送り致しますので、お名前とご住所を、お聞かせ願えますか？

3002|1400
1500@I will then, also enclose a presentation application form for you.
1600@Please fill in the necessary information and [send] it back [to] us.
1400@発表申込み用紙を、同封致しますので、それに必要事項を記入して、まず送って下さい。

3003|700
800@Okay, then we will [send] the registration forms and the presentation application form [to] you.
900@Would you please fill in the necessary information and [send] it back [to] us?
700@そうですか。
800@それでは、こちらの方から、登録用紙と発表申込みの為の書類を、御送りしますので、そこに必要事項を記入して戴き、こちらの方へお送り願えますでしょうか。

3003|1200
1500@Then I will [send] five registration forms and one presentation application [to] you.
1600@If you would fill them in and [send] them all back [to] us together, that would be fine.
1500@では、こちらから登録用紙５名分と、発表申込み書１名分御送り致しますので、それらをまとめて、返送して戴ければ結構です。

3003|1500
1900@Okay, then I will [send] the presentation and registration forms [to] this address.
1800@はい、承知致しました。
1900@それではそちらの住所の方に、登録用紙及び発表書類を、御送り致します。

3003|1600
2000@It is okay to put them all together and [send] them [to] this address, isn't it?
2000@全部まとめて、今戴いた御住所の方へ、送らせてもらってよろしいですね。

以上

(03)　　8件 形態素情報
　　係り受け情報

****
検索依頼ドキュメント
****

1．依 頼 者： 竹澤 研究員

2．依頼日時： 91/10/09

3．回答日時： 91/10/16

4．作業日数： 5日間

5．回 答 者： 井上 健吾 （株）東洋情報システム

6．内　　容： 係受け情報の出力

7．対象会話： ＡＤＤ登録済みデータ全部

8．出力形式： 構文関係名__意味関係名__000__:_係り元文節文字列:_¥
係り元標準表現__:_XXXXXXX__;_____;_係り先文節文字列¥
:_係り先標準表現__:_XXXXXXX
（"_"はスペース）

9．出力サンプル： 下欄に別記する。

10．検索手順： 検索別バッチファイル参照 （別冊添付資料）

11．ファイル： 使用ファイルの一覧
($SEARCH = /data6/LDB/search, $TBMT = /data6/LDB/TBMT/KB)

検索結果:
$SEARCH/ans/TAKEZAWA/TAKEZAWA3/kakari.data

バッチプログラム:
$SEARCH/work/TAKEZAWA/TAKEZAWA3/TAKEZAWA3.bat

データソース:
ＡＤＤを使用

データ加工用ツール:
$SEARCH/tool/kifilter

$TBMT/tool/REFORM
/JOIN2
/RECONSTRUCT
/CKSORT

12．内容補足：文節文字列は単語単位に区切り"-"が入っています。

13．備　　考：なし

14．依頼ＩＤ：TAKEZAWA3

****

#出力サンプル

ファイル：
atr-dp:/data6/LDB/search/ans/TAKEZAWA/TAKEZAWA3/kakari.data

出力形式：
構文関係名__意味関係名__000__:_係り元文節文字列:_係り元標準表現__:_XXXXXXX¥
__;_____:_係り先文節文字列:_係り先標準表現__:_XXXXXXX
("_"はスペース)

検索結果：
02 AVO 000 : 通訳電話国際会議-の: 通訳電話国際会議 : XXXXXXX ; : 第１回 の: 回 : XXXXXXX
01 OBJ 000 : 事務局-でしょ-う-か: です : XXXXXXX ; : そちら: そちら : XXXXXXX
01 PRD 000 : 事務局-でしょ-う-か: です : XXXXXXX ; : 事務局 でしょ う か: 局 : XXXXXXX
02 PRP 000 : 事務局-でしょ-う-か: 局 : XXXXXXX ; : 通訳電話国際会議 の: 通訳電話国際会議 : XXXXXXX
02 RNG 000 : こと-で-ね: 事 : XXXXXXX ; : 会議 の: 会議 : XXXXXXX
02 RNG 000 : こと-で: 事 : XXXXXXX ; : 登録 の: 登録 : XXXXXXX
01 CON 000 : おうかがい-し-たい-ん-です-が: たい : XXXXXXX ; : おうかがい し たい ん です が: うかがい : XXXXXXX
01 RNG 000 : おうかがい-し-たい-ん-です-が: 伺う : XXXXXXX ; : こと で ね: こと : XXXXXXX
01 RNG 000 : おうかがい-し-たい-ん-です-が: 伺う : XXXXXXX ; : こと で: こと : XXXXXXX
04 DGR 000 : おうかがい-し-たい-ん-です-が: たい : XXXXXXX ; : ちょっと: ちょっと : XXXXXXX
01 OBJ 000 : ある-ん-です-けれども: 有る : XXXXXXX ; : 登伺用紙 が: 用紙 : XXXXXXX
01 SPA 000 : ある-ん-です-けれども: 有る : XXXXXXX ; : 手元 に: 手元 : XXXXXXX
01 TMA 000 : ある-ん-です-けれども: 有る : XXXXXXX ; : 今: 今 : XXXXXXX
02 DLM 000 : 中-で: 中 : XXXXXXX ; : その: その : XXXXXXX
02 OAT 000 : 名前-と: 名前 : XXXXXXX ; : クレジットカード の: クレジットカード : XXXXXXX
02 DLM 000 : ナンバー-を: ナンバー : XXXXXXX ; : なん か: なん : XXXXXXX
01 OBJ 000 : 書く: 書く : XXXXXXX ; : クレジットカード を ね: クレジットカード : XXXXXXX
01 OBJ 000 : 書く: 書く : XXXXXXX ; : ナンバー を: ナンバー : XXXXXXX
01 OBJ 000 : 書く: 書く : XXXXXXX ; : 名前 と: 名前 : XXXXXXX
01 SPA 000 : 書く: 書く : XXXXXXX ; : 中 で: 中 : XXXXXXX

****
検索依頼ドキュメント
****

１．依 頼 者： 浦谷 主幹研究員

２．依頼日時： 91/10/18

３．回答日時： 91/10/30

４．作業日数： ８日間

５．回 答 者： 井上 健吾 （株）東洋情報システム

６．内　　容： 条件に合った係り受けの係り先、係り元の出力
（係り元の文節も出力）

条件：
1. 構文コードが"04","05"のもの。
2. 構文コードが"01"でかつ意味コードが"CON","DGR"のもの
3. 1,2の条件に合うもので係り先、係り元の単語が共に以下の品詞のもの。
動詞、サ変名詞、形容名詞、形容詞

７．対象会話： 領　　域: 国際会議
メディア: 電話/キーボード

８．出力形式： 係り先単語（標準表現）|係り元単語（標準表現）|¥
係り元文節（標準表現）|接続詞の有無

９．出力サンプル： 下欄に別記する。

１０．検索手順： 検索別バッチプログラム参照（別冊添付資料）

１１．ファイル： 使用ファイルの一覧
($SEARCH = /data/LDB/search, $TBMT = /data6/LDB/TBMT/KB)

検索結果:
$SEARCH/ans/URATANI/URATANI2/jrel.ans

バッチプログラム:
$SEARCH/work/URATANI2.bat

データソース:
$SEARCH/work/TAKEZAWA/TAKEZAWA3/jword.mast

データ加工用ツール:
$SEARCH/tool/pickup_data
/kifilter

$TBMT/tool/REFORM
/JOIN2
/RECONSTRUCT

１２．内容補足： 係り元文節は係り元単語を"～"に変換しています.

１３．備　　考：なし

１４．依頼ＩＤ：URATANI2

****

#出力サンプル

ファイル名：
atr-dp:/data6/LDB/search/ans/URATANI/URATANI2/jrel.ans

出力形式：
係り先単語（標準表現）|係り元単語（標準表現）|係り元文節（標準表現）|接続詞の有無

検索結果：
有る|有る|～んですけれども|有
結構|持つ|御～だないば|有
成る|有る|～んですが|有
成る|行く|～ますと|有
結構|成る|～ますので|有
行ける|申込む|～ないば|有
速い|参加|～するんですけれども|有
速い|言う|～て|有
行く|行く|～て|有
成る|来る|～られるますと|有
成る|参る|～ますけれども|有
成る|高い|～|無
思う|無い|～と|無
成る|短い|～|無
掛る|出る|～ておるますて|有
思う|行く|～うと|無
着く|乗る|御～になるば|有
成る|乗る|御～になるば|有
成る|着く|～ますし|有
無い|成る|～ているますたが|有
成る|言う|～ますても|有
思う|成る|～と|無
成る|成る|～ているますので|有
宜しい|送る|御～するば|有
知れる|無い|～かも|無
思う|知れる|～ないと|無
思う|出来る|～ないんですが|有
行ける|間に合う|～ますが|有
遣る|急ぐ|～で|有
思う|行ける|～ないと|無
延ばす|変更|～するますて|有
成る|変更|～するますて|有
宜しい|送る|御～するば|有
結構|教える|御～いたすますけれども|有
提出|急ぐ|～で|有
結構|入れる|～ますので|有
成る|使える|～ようだ|無
願う|送る|御～するますので|有
聞く|有る|～と|無
思う|聞く|～たんですけれども|有
思う|掛る|～ておるますて|有
思う|分る|～ますんので|有
話す|分る|～ようだ|無
話す|かい摘む|～で|有
総括的|非常|～だ|無
専門的|非常|～だ|無
思う|言える|～と|無
おもしろい|非常|～だ|無
思う|おもしろい|～のだはないかと|無
思う|出る|～てまいるますけれども|有

(04)　　4件 原データのみ

****
検索依頼ドキュメント
****

1．依 頼 者： 山岡 研究員

2．依頼日時： 90/11/05

3．回答日時： 90/11/05

4．作業日数： 5 日間

5．回 答 者： 井上・庄山 （株）東洋情報システム

6．内　　容： 検索パターンを含む発話とその前発話を出力
パターン:
「はい」「いいえ」「そうです」「そうですか」
「わかりました」「お願いします」
（それぞれ"、"と"。"を付ける２つのパターンで検索しました）

7．対象会話： 領　　域: 国際会議
メディア: 電話

8．出力形式： #
sp1: 日本語文　（前発話）
日本語文
日本語文
sp2: 日本語文　（検索パターンを含む発話）
日本語文
日本語文
#　← ２組の発話の区切りとして"#"を挿入します

9．出力サンプル： 下欄に別記する。

10．検索手順： 下欄に別記する。

11．ファイル： 使用ファイルの一覧 ($SEARCH = /data/LDB/search)

検索結果:
$SEARCH/ans/YAMAOKA/YAMAOKA1/utter_and_prev_utter.ans
/cnt #会話数等を出力

個別プログラム:
$SEARCH/work/YAMAOKA/YAMAOKA1/YAMAOKA1.bat

データソース:
/data3/MORPH/bunout/key/*　#文節情報ファイル

データ加工用ツール:
$SEARCH/tool/jwrd_to_sent

12．内容補足： 検索対象の会話数、発話数、文数、文節数、単語数も提出しています

13．備　　考：

14．依頼ＩＤ：YAMAOKA1

****
次ページへ続く

＃出力サンプル

・utter_and_prev_utter.ans

```
  #
  sp1: 日本語文　（前発話）
       日本語文
       日本語文
  sp2: 日本語文　（後発話）
       日本語文
       日本語文
  #

sp1 : はい、参加予定者は、全部で五名います。
      学生が三名と研究生が一人、この者が発表をします。
      それと教授の私です。
      まとめて手続出来ますか？
sp2 : はい。
      お出来になります。
      では、こちらから登録用紙５名分と、発表申込み書１名分御送り致しますので、それらをまとめ
て、返送して戴ければ結構です。
#
sp1 : もしもし。
sp2 : はい。
#
sp1 : もしもし。
sp2 : はい。
      こちら通訳会議電話事務局ですが。
#
```

・ cnt

```
   会話数| 発話数| 文数| 文節数| 単語数

Sum total:  175files|      2201|      4769|     24132|     69826
```

＃検索手順

```
1.
# 処理内容：文節情報ファイルより検索に必要な情報を抽出する
# 入力形式：文節情報ファイル
# 出力形式：会話ＩＤ|発話ＩＤ|文ＩＤ|出現単語|標準表現|品詞コード

2.
# 処理内容：会話ＩＤ、発話ＩＤ、文ＩＤをＫＥＹとして
#           出現単語を文単位の文字列にします
# 入力形式：会話ＩＤ|発話ＩＤ|文ＩＤ|出現単語|標準表現|品詞コード
# 出力形式：会話ＩＤ|発話ＩＤ|出現単語出現単語出現単語．．．（文単位）

3.
# 処理内容：検索パターンを含む日本語文を抜き出す
# 入力形式：会話ＩＤ|発話ＩＤ|出現単語出現単語出現単語．．．（文単位）
# 出力形式：会話ＩＤ|発話ＩＤ|出現単語出現単語出現単語．．．（文単位）

4.
# 処理内容：検索パターンを含む文を含む発話をその発話の前発話と共に出力する
# 入力形式：会話ＩＤ|発話ＩＤ|出現単語出現単語出現単語．．．（文単位）
# 出力形式：#
#           sp1: 出現単語出現単語出現単語．．．（文単位） 前発話
#                出現単語出現単語出現単語．．．（文単位）
#           sp2: 出現単語出現単語出現単語．．．（文単位） 後発話
#                出現単語出現単語出現単語．．．（文単位）
#           # ← ２組の発話ごとに区切りとして挿入
```

以上

(05)　　３件 形態素情報
　　日英対応情報:発話対応

****
検索依頼ドキュメント
****

１．依 頼 者： 隅田 研究員

２．依頼日時： 91/05/09

３．回答日時： 91/05/22

４．作業日数： １０日間

５．回 答 者： 井上 健吾 （株）東洋情報システム

６．内　　容： 発話単位の日英対応を１ｉｓｐ形式で出力

７．対象会話： 領　　域: 国際会議
メディア: 電話
言語パターン: 日英／英日／英日（同時通訳）

８．出力形式： ("発話対応ＩＤ-英語発話者-日本語発話者"
"英語発話対応文字列"
"日本語発話対応文字列"
<RET>
<RET>
)

９．出力サンプル： 下欄に別記する。

１０．検索手順： 検索別バッチプログラム参照 （別冊添付資料）

１１．ファイル： 使用ファイルの一覧
($SEARCH = /data/LDB/search, $TBMT = /data6/LDB/TBMT/KB)

検索結果:
$SEARCH/ans/SUMITA/SUMITA4/interpret/questioner/jeut.*
/secretariat/jeut.*
エラーデータ:
/interpret-error/questioner/jeut.*
/secretariat/jeut.*

バッチプログラム:
$SEARCH/work/SUMITA/SUMITA4/SUMITA4.bat

データソース:
ＳＱＬにより検索

データ加工用ツール:
$SEARCH/tool/kifilter2
/catline3

$TBMT/tool/CKSORT
/JOIN
/RECONSTRUCT

１２．内容補足：検索結果は会話番号別にファイルを分割してあります。

１３．備　　考：発話単位での対応を出力しているが、日本語発話者と英語発話者が異なっていなければデータとし

て不良である。　この様なデータが今回の検索で
発見されたため、此のデータはエラーデータとし
て他のデータとは別けて出力した。

１４．依頼ＩＤ：SUMITA4

****

#出力サンプル

ファイル:
atr-dp:/data6/LDB/search/ans/SUMITA/SUMITA4/interpret/questioner/jeut.101

出力形式:
("発話対応ＩＤ-英語発話者-日本語発話者"
<TAB>"英語発話対応文字列"
<TAB>"日本語発話対応文字列"
<RET>
<RET>
)

検索出力:
("0010-質問者-通訳者"
"Good morning .
This is Professor Chris Michelson from the University of Toronto .
I'm trying to reach Mr. Hiroshi Okamoto , please ."
"トロント大学のクリス・マイケルセンですけれども、岡本宏さんお願いしたいと思います。"

)
("0020-通訳者-事務局"
"This is Okamoto speaking at the Secretariat of the ATR ."
"はい、もしもし、ＡＴＲ事務局の岡本でございます。"

)
("0030-質問者-通訳者"
"Good morning , Mr. Okamoto .
I'm not sure if you remember me , I'm a member of the Planning Committee for the International Conference on Computer Science .
I'm calling regarding my speech , my opening greetings to be delivered on the 28th < twenty-eighth > , I believe that's a Friday .
I'm afraid I've been extremely busy , I've have certainly given time to the speech , but I haven't had a chance to quite finish it .
You requested a typed speech for Monday to be sent to you on Monday , the day after , well , Monday .
I'm afraid I won't have typed copy ready , I will have a handwritten copy available .
I must attend a meeting today and will have a chance to work on the speech this afternoon .
I will be able to send you a typed copy next Thursday or Friday .
I hope this won't be a problem ."
"岡本さんでいらっしゃいますか。
覚えていらっしゃいますでしょうか。
国際コンピューター会議の企画委員の一人でございます。
で、今日は２８日の金曜日に私が発表することになっておりますのでその挨拶文について少しお伺いしたいと思いますけれども、非常に時間がなくて忙しい毎日を送っておりますので、なかなか出来上がるのが遅くなっているんですが、月曜日までにタイプアップをしたスピーチを送るということでしたけれども、タイプアップちょっと今すぐにできませんので、手書きであれば今日会議に出席しました後、午後仕上げまして、お送りすることができます。
タイプアップしたものについては今度の木曜日か金曜日であればお送りできると思うんですが。"

)

以上

(06)　　2件 形態素情報
　　日英対応情報:単語対応
　　係り受け情報

****
検索依頼ドキュメント
****

1．依 頼 者：井ノ上研究員

2．依頼日時：90/06/01

3．回答日時：90/06/05

4．回 答 者：高橋・庄山(TIS)

5．内　　容： 係受けおよび対訳情報の検索.

6．対象会話： 全会話

7．出力形式： 会話ＩＤ,被修飾語(HEAD),読み,品詞,対訳,被修飾語(MODIFIER),読み,品詞,対訳,構文コード,意味コード

8．検索手順：下欄に別記する.

9．ファイル：($SEARCH → /data6/LDB/search)
　　ldbshパターンファイル
　　$SEARCH/work/INOUE/INOUE3/EJ-kakari.pat

　　ldbsh出力
　　$SEARCH/work/INOUE/INOUE3/EJ-kakari.ans

　　フォーマット変換プログラム
　　$SEARCH/work/INOUE/INOUE3/EJ-kakari.rew

　　検索結果
　　$SEARCH/work/INOUE/INOUE3/EJ-kakari.out

１０．内容補足：修飾語、被修飾語はこれを含む文節の形で出す。
　　読み、品詞はなくても構わない。

１１．備　　考：INOUE1を全会話に対して行なった.

１２．依頼ＩＤ：INOUE3

##########

検索手順

1．editpでパターンファイルを作成する。

```
editp EJ-kakari.pat
```

2．psearchで検索実行。

```
psearch EJ-kakari.pat > EJ-kakari.ans
```

3．EJ-kakari.rewでフォーマット変換。

```
EJ-kakari.rew EJ-kakari.ans > EJ-kakari.out
```

以上。

(07)　　2件 形態素情報
日英対応情報:単語対応,文対応

****
検索依頼ドキュメント
****

1．依 頼 者： 長谷川 研究員

2．依頼日時： 90/10/01

3．回答日時： 90/10/04

4．作業日数： 4日間

5．回 答 者： 井上・庄山 （株）東洋情報システム

6．内　　容： 検索対象単語とその対訳を含む日英の文を出力
会話ＩＤ、単語ＩＤをＫＥＹとしてソートしたものと、
日本語単語をＫＥＹとしてソートした出力結果があります。

検索対象単語: （標準表現）
有る 思う 申す 成る 為る 御座います 出来る 教える 願う 分かる
聞く 行く 持つ 待つ 伺う 支払う 掛る 言う 載く 話す 考える
書く 見る 会う 行う 載ける

7．対象会話： 領　　域: 国際会議／旅行
メディア: 電話／キーボード

8．出力形式： 会話ＩＤ|単語ＩＤ|日本語単語|英語単語|日本語文|英語文

9．出力サンプル： 下欄に別記する。

10．検索手順： 下欄に別記する。

11．ファイル： 使用ファイルの一覧 ($SEARCH = /data/LDB/search)

検索結果:
$SEARCH/ans/HASEGAWA/HASEGAWA1/HASEGAWA1.ans # 会話ＩＤ,単語ＩＤをＫＥＹとしてソート
/HASEGAWA1.ans2 # 日本語単語をＫＥＹとしてソート

個別プログラム:
$SEARCH/work/HASEGAWA1.bat

データソース:
$SEARCH/lib/JWORD-regu.all
/EWORD.all
/j_word.all （このファイルは現在存在しません）
/e_word.all （このファイルは現在存在しません）

データ加工用ツール:
$SEARCH/tool/jwrd_to_sent
/ewrd_to_sent

12．内容補足： 検索対象会話は９０年６月に存在した全ファイル

13．備　　考：

14．依頼ＩＤ：HASEGAWA1

****
次ページへ続く

＃出力サンプル
・HASEGAWA1.ans （会話ＩＤ、単語ＩＤをＫＥＹとしてソート）
会話ＩＤ|単語ＩＤ|日本語単語|英語単語|日本語文|英語文

41|34500|なっ|is|郵便番号が１００の、２６となっております。|And the address , the Zip code is 100 . 100 26
41|45400|書い|check|そうしましたら、まず、コールフォーペーパース、お送りしますので、それにすべて、費用ですとか、詳しい事、書いてありますので、そちらの方、参照していただきたいと思うんですが。|We will send you our call-for paper as soon as possible . we think you can check a detailed information from the paper . Would you refer to that ?
41|52000|なっ|is|はい、登録費用は、１００ドルになってます。|Registration fee is 100 U.S. dollars .

・HASEGAWA1.ans2 （日本語単語をＫＥＹとしてソート）
会話ＩＤ|単語ＩＤ|日本語単語|英語単語|日本語文|英語文

169|91700|あっ|be|あとそうですね、最後の日も分科会があって、この記念講演というので一つの会場の方に戻るわけですね。|On the last day , there'll be sessions , and we'll have to go back to the hall for the memorial speech .
175|48800|あっ|expect|もしもし、この間、去年だったかなあ、エーアイーのセミナーが東京であったんだけども、そのときがたしか正会員じゃなくて一般の登録料が九万円で、それからそのとき学会の割引ということで六万円の登録料だったんだけども、当然同じエーアイなんだから、そういう割引があってもいいんじゃないんでしょうかね。|Maybe last year , at the AI seminar in Tokyo , the ordinary registration fee was 90,000 yen and the special fee was 60,000 yen for members of the societies . This conference is also in the field of AI , and I think we can expect the same kind of consideration .

＃検索手順

1.
処理内容：/data6/LDB/search/lib/j_word.all(e_word.all)より検索に必要なデータを抜き出す。
入力形式：/data6/LDB/search/lib/j_word.all(e_word.all)
出力形式：日: 会話ＩＤ-単語対応|単語ＩＤ|出現単語|標準表現
#　　英: 会話ＩＤ@単語対応|文ＩＤ|出現単語

2.
処理内容：会話ＩＤ、単語対応をＫＥＹとして出現単語を単語対応単位の文字列とする。
入力形式：英: 会話ＩＤ@単語対応|文ＩＤ|出現単語
出力形式：英: 会話ＩＤ@単語対応@文ＩＤ|出現単語 出現単語 出現単語 ...

3.
処理内容：日本語単語データより検索対象単語を抜く出す。
入力形式：会話ＩＤ-単語対応|単語ＩＤ|出現単語|標準表現
出力形式：会話ＩＤ-単語対応|単語ＩＤ|出現単語|標準表現（検索対象単語）

4.
処理内容：会話ＩＤ、単語対応をＫＥＹとして日英の対訳単語を連接状態にする。
入力形式：日: 会話ＩＤ-単語対応|単語ＩＤ|出現単語|標準表現
#　　英: 会話ＩＤ@単語対応@文ＩＤ|出現単語 出現単語 出現単語 ...
出力形式：会話ＩＤ-文ＩＤ|単語:単語ＩＤ:英単語

5.
処理内容：会話ＩＤ、文対応をＫＥＹとして日本語単語を文単位の文字列とします。
入力形式：/data6/LDB/search/lib/JWORD-regu.allを使用しています
出力形式：会話ＩＤ-文対応ＩＤ|日本語文

6.
処理内容：会話ＩＤ、文対応をＫＥＹとして英語単語を文単位の文字列とします。
入力形式：/data6/LDB/search/lib/EWORD.allを使用しています
出力形式：会話ＩＤ-文対応ＩＤ|日本語文

7.
処理内容：入力形式に文対応ＩＤを入れフォーマットを変更します
入力形式：会話ＩＤ-文ＩＤ|単語:単語ＩＤ:英単語
出力形式：会話ＩＤ-文対応|単語:単語ＩＤ:英単語
コメント：文対応ＩＤは/data6/LDB/search/lib/EWORD.allを参照します

8.
処理内容："会話ＩＤ-文対応ＩＤ"をＫＥＹとして入力形式に日英の文を追加します
入力形式：会話ＩＤ-文対応|単語:単語ＩＤ:英単語
出力形式：会話ＩＤ-文対応|単語ＩＤ:単語:英単語@日文|英文

9.
処理内容：提出用のフォーマットに変更しソートします
入力形式：会話ＩＤ-文対応|単語ＩＤ:単語:英単語@日文|英文
出力形式：会話ＩＤ|単語ＩＤ|日本語単語|英語単語|日文|英文
コメント：ソートは"会話ＩＤ|単語ＩＤ"をＫＥＹにしたもの、
#　　"日本語単語"をＫＥＹにしたものの２つがあります

以上

(08)　　1件 形態素情報
　　日英対応情報:文節対応
　　係り受け情報

****
検索依頼ドキュメント
****

1．依 頼 者： 江原 主幹研究員

2．依頼日時： 91/02/05

3．回答日時： 91/04/01

4．作業日数： 9日間

5．回 答 者： 井上 健吾 （株）東洋情報システム

6．内　　容： 日本語文節間の係り受けデータとそれに対応する日英文節
対応データを出力。

7．対象会話： 領　　域: 国際会議
メディア: キーボード

8．出力形式： 会話ＩＤ|MODIFIER文節ＩＤ|HEAD文節ＩＤ|MODIFIER文節文字列|¥
HEAD文節文字列|MODIFIER文節対応|HEAD文節対応|¥
MODIFIER英語文節文字列|HEAD英語文節文字列

9．出力サンプル： 下欄に別記する。

10．検索手順： 検索別バッチプログラム参照 （別冊添付資料）

11．ファイル： 使用ファイルの一覧
($SEARCH = /data/LDB/search, $TBMT = /data6/LDB/TBMT/KB))

検索結果:
$SEARCH/ans/EHARA/EHARA5/kakariuke.jp_ep

バッチプログラム:
$SEARCH/work/EHARA/EHARA5/EHARA5.bat

データソース:
ＡＤＤを使用

データ加工用ツール:
$SEARCH/tool/pickup_data

$TBMT/tool/JOIN
/RECONSTRUCT
/CKSORT

12．内容補足： なし

13．備　　考： ４／１に提出した結果にバグがあったため４／５に再提出
しています。

14．依頼ＩＤ：EHARA5

****

＃出力サンプル

ファイル:
atr-dp:/data6/LDB/search/ans/EHARA/EHARA5/kakariuke.jp_ep

出力形式:
会話ＩＤ|MODIFIER文節ＩＤ|HEAD文節ＩＤ|MODIFIER文節文字列|HEAD文節文字列|¥
MODIFIER文節対応|HEAD文節対応|MODIFIER英語文節文字列|HEAD英語文節文字列

検索結果:
3001|500|600|会議に|申込みたいのですが。|200|300|for the conference|I would like to apply
3001|800|1000|登録用紙は|お持ちでしょうか？|400|600|the application forms|do you ## have
3001|900|1000|既に|お持ちでしょうか？|500|600|already|do you ## have
3001|1500|1600|こちらから|お送り致しますので、|900|1000|I|will send them to you
3001|1600|1900|お送り致しますので、|お聞かせ願えますか？|1000|1300|will send them to you|Would you please give me
3001|1700|1900|お名前と|お聞かせ願えますか？|1100|1300|your name and|Would you please give me
3001|1800|1900|ご住所を、|お聞かせ願えますか？|1200|1300|address|Would you please give me
3001|3000|3100|至急に|送らせていただきます。|2100|2200|immediately|I'll send the forms out
3001|3300|3500|何か|点が、|2300||anything|
3001|3500|3400|点が、|分からない||2400||you don't understand

以上

(09)　　1件 形態素情報
日英対応情報:文節対応

****
検索依頼ドキュメント
****

1．依 頼 者： 江原 主幹研究員

2．依頼日時： 90/11/15

3．回答日時： 90/11/26

4．作業日数： 4 日間

5．回 答 者： 井上・庄山 （株）東洋情報システム

6．内　　容： 文節対応のデータを見やすく出力（日英それぞれ出力）

7．対象会話： atr-dp:/data6/XLDB/data/key/ID-SET にある会話ＩＤ

8．出力形式： 1. 日本語文節対応
会話ＩＤ|文ＩＤ|文節ＩＤ|文節対応|文節単位の文字列

2. 英語文節対応
会話ＩＤ|文ＩＤ|文節対応|文節単位の文字列

9．出力サンプル： 下欄に別記する。

10．検索手順： 下欄に別記する。

11．ファイル： 使用ファイルの一覧 ($SEARCH = /data/LDB/search)

検索結果:
$SEARCH/ans/EHARA/EHARA4/jphrase.ans
/ephrase.ans

個別プログラム:
$SEARCH/work/EHARA/EHARA4/EHARA4.bat

データソース:
$SEARCH/lib/TABLES/j_word.idx
/j_paharse.idsx
/e_word.idx

データ加工用ツール:
$SEARCH/tool/select-from.csh

12．内容補足： 文対応ＩＤが"-1"の場合は文対応ＩＤは空欄としました

13．備　　考：

14．依頼ＩＤ：EHARA4

****
次ページに続く

#出力サンプル
・jpharse.ans
会話ＩＤ|文ＩＤ|文節ＩＤ|文節対応|文節単位の文字列

```
3045|100|100|100|もしもし、
3045|100|200|200|来る
3045|100|300|300|コンピューター会議に付いて
3045|100|400|400|質問したいのですが。
3045|200|500|500|はい、
3045|200|600|600|こちら
3045|200|700|700|事務局ですが、
3045|200|800|800|どうぞ。
3045|300|900|900|こんにちは。
3045|400|1000|1000|自己紹介します。
```

・ephrase.ans
会話ＩＤ|文ＩＤ|文節対応|文節単位の文字列

```
3045|100|100|Hello
3045|100||,
3045|100|400|I would like to make an inquiry
3045|100|300|on the
3045|100|200|upcoming
3045|100|300|computer conference
3045|100||.
3045|200|500|Yes
3045|200||,
3045|200|600|this
```

#検索手順

```
1.
# 処理内容：$TABLEPATH/j_word.idxより"会話ＩＤ"、"文節ＩＤ"、"出現単語"
#           を抽出。
# 入力形式：/data6/LDB/search/lib/TABLES/j_word.idx を使用
# 出力形式：会話ＩＤ@文節ＩＤ|出現単語 （ファイル名:w_jwrd）

2.
# 処理内容：$TABLEPATH/j_phrase.idxより"会話ＩＤ"、"文節ＩＤ"、"文ＩＤ"、
#           "文節対応" を抽出。
# 入力形式：/data6/LDB/search/lib/TABLES/j_phrase.idx を使用
# 出力形式：会話ＩＤ@文節ＩＤ|文ＩＤ@文節対応 （ファイル名:w_jph）

3.
# 処理内容："会話ＩＤ-文節ＩＤ"をＫＥＹとしてw_jwrdとw_jphを連接状態にする。
# 入力形式：w_jwrd: 会話ＩＤ@文節ＩＤ|出現単語
#           w_jph : 会話ＩＤ@文節ＩＤ|文ＩＤ@文節対応
# 出力形式：会話ＩＤ@文ＩＤ@文節ＩＤ|文節対応|出現単語（ファイル名:jword.list）

4.
# 処理内容：$TABLEPATH/e_word.idxより"会話ＩＤ"、"文ＩＤ"、"出現単語"、
#           "文節対応"を抽出。
# 入力形式：/data6/LDB/search/lib/TABLES/e_word.idx を使用
# 出力形式：会話ＩＤ@文ＩＤ@文節対応|出現単語 （ファイル名:eword.list）

5.
# 処理内容："会話ＩＤ"、"文ＩＤ"、"文節ＩＤ"、"文節対応"をＫＥＹとして
#           日本語出現単語を連接状態にします。
# 入力形式：会話ＩＤ@文ＩＤ@文節ＩＤ|文節対応|出現単語（ファイル名:jword.list）
# 出力形式：会話ＩＤ|文ＩＤ|文節ＩＤ|文節対応|出現単語出現単語．．．
#                                        （ファイル名:jphrase.ans）

6.
# 処理内容："会話ＩＤ"、"文ＩＤ"、"文節対応"をＫＥＹとして日本語出現単語を
#           連接状態にします。
```

```
# 入力形式：会話ＩＤ|文ＩＤ|文節対応|出現単語（ファイル名:eword.list）
# 出力形式：会話ＩＤ|文ＩＤ|文節対応|出現単語出現単語出現単語．．．
#                                        （ファイル名:ephrase.ans）
```

以上

(10)　　1件 形態素情報
　　日英対応情報:単語対応

****
検索依頼ドキュメント
****

1．依 頼 者：井ノ上研究員

2．依頼日時：90/04/24

3．回答日時：90/04/26

4．回 答 者：高橋・庄山(TIS)

5．内　　容：対訳データ

6．対象会話： 全会話

7．出力形式： 標準表現|訳語|頻度

8．検索手順：下欄に別記する.

9．ファイル：($SEARCH → /data6/LDB/search)
　　検索用バッチファイル
　　$SEARCH/work/INOUE/INOUE2/INOUE2.bat

　　検索プログラム
　　$SEARCH/work/INOUE/INOUE2/ewrd-wcorr2
　　id_to_ewrd
　　id_to_jwrd.reg.pos
　　検索結果
　　$SEARCH/work/INOUE/INOUE2/taiyaku.koyuumeisi
　　taiyaku.daimeisi
　　taiyaku.sahenmeisi
　　taiyaku.futsuumeisi
　　taiyaku.hondousi

10．内容補足：

11．備　　考：

12．依頼ＩＤ：INOUE2

##########

検索手順

1．

```
### PROCESURE 1 ##########################################################
# 処理内容：w_corrより、対応する日本語単語ＩＤ、英語単語ＩＤおよび会話ＩＤを  #
#           検索出力する。                                                #
# 入力形式：会話ＩＤ,会話ＩＤ,会話ＩＤ,・・・
# 出力形式：会話ＩＤ|日本語単語ＩＤ|英語単語ＩＤ
##########################################################################
```

2．

```
### PROCESURE 2 ##########################################################
# 処理内容：日本語単語ＩＤ、英語単語ＩＤおよび会話ＩＤをもとに、相対応する日英#
#           の単語を検索出力する。                                        #
# 入力形式：会話ＩＤ|日本語単語ＩＤ|英語単語ＩＤ                          #
# 出力形式：日本語・標準表現|日本語・品詞|英語・出現単語                  #
# 注　　意：日英の単語の検索にそれぞれ id_to_jwrd.reg, id_to_ewrdを使用。  #
#           while ループでの時間短縮のため入力ファイルを分割しています。  #
#           分割サイズはSPLIT_SIZEに設定してあります。                    #
##########################################################################
```

3．

```
### PROCESURE 3 ##########################################################
# 処理内容：件数のカウント
# 入力形式：日本語・標準表現|日本語・品詞|英語・出現単語                  #
# 出力形式：日本語・標準表現|日本語・品詞|英語・出現単語|件数              #
##########################################################################
```

4．

```
### PROCESURE 4 ##########################################################
# 処理内容：目的の品詞のデータを取りだして、フォーマット出力。
# 入力形式：日本語・標準表現|日本語・品詞|英語・出現単語|件数
# 出力形式：日本語・標準表現|英語・出現単語|件数
##########################################################################

# 以上。
```

(11)　　1件 形態素情報
　　係り受け情報
　　日英対応情報:格対応

****
検索依頼ドキュメント
****

1．依 頼 者： 江原 主幹研究員

2．依頼日時： 90/07/19

3．回答日時： 90/08/01

4．回 答 者： 高橋.庄山 (TIS)

5．内　　容： 格構造対応について、対応有りの度数と対応なしの度数

6．対象会話： 指定の会話ＩＤ(→9.を参照)

7．出力形式： TYPE1 件数|TYPE2 件数|TYPE3 件数
TYPE1: 完全対応(すべての格要素が対応している)
TYPE2: 不完全対応(すべての格要素が対応している)
TYPE3: 対応なし(どの格要素も対応していない)

8．検索手順： 下欄に別記する。

9．ファイル： 使用ファイルの一覧 ($SEARCH = /data/LDB/search)

検索結果:
$SEARCH/ans/EHARA/EHARA1/JE-kaku2.tel
JE-kaku2.key

対象会話:
$SEARCH/work/EHARA/EHARA1/ID-set.tel
ID-set.key

個別プログラム:
$SEARCH/work/EHARA/EHARA1/je-kaku2
work/EHARA/EHARA1/kakutaiou-check.csh
work/EHARA/EHARA1/and.csh
tool/kakufilter.csh

データ加工用ツール:
$SEARCH/tool/jwrd_to_sent
ewrd_to_sent
catline3

10．内容補足： 他の格構造を含んでいる格構造は調査の対象としない.

11．備　　考：/data6/XLDB/pgm/je-kakuの修正版

12．依頼ＩＤ：EHARA1

****

検索手順

```
### PROCESURE 1 ##############################################################
# 処理内容：格対応テーブルの検索
# 入力形式：会話ＩＤリスト
# 出力形式：会話ＩＤ|構造ＩＤ|単語ＩＤ(擬似マルチ)|格対応
##############################################################################
### PROCESURE 2 ##############################################################
# 処理内容：同じ構造ＩＤを持つ単語ＩＤを1レコードにする.
# 入力形式：会話ＩＤ|構造ＩＤ|単語ＩＤ擬似マルチ|格対応
# 出力形式：会話ＩＤ-構造ＩＤ|単語ＩＤ擬似マルチ(会話ＩＤと構造ＩＤがユニーク)
##############################################################################
### PROCESURE 3 ##############################################################
# 処理内容：1レコード内の単語ＩＤをソートする.
#           先頭と末尾以外の単語ＩＤを削除する.
# 入力形式：会話ＩＤ-構造ＩＤ|単語ＩＤ擬似マルチ(会話ＩＤと構造ＩＤがユニーク)
# 出力形式：会話ＩＤ-構造ＩＤ|先頭の単語ＩＤ|末尾の単語ＩＤ
##############################################################################
### PROCESURE 4 ##############################################################
# 処理内容：連続する２つの格構造レコードに着目して、前後の包含関係を調べる.
#           他の各構造を含んでいる場合は、何もしない.
#           他の各構造を含んでいない場合は、前の構造ＩＤを出力する.
# 入力形式：会話ＩＤ-構造ＩＤ|先頭の単語ＩＤ|末尾の単語ＩＤ
# 出力形式：会話ＩＤ-構造ＩＤ
##############################################################################
### PROCESURE 5 ##############################################################
# 処理内容：構造ＩＤ毎に格要素の対応率を算出する
# 入力形式："会話ＩＤ"|"構造ＩＤ"|"格対応ＩＤ"(格要素毎に１レコード)
# 出力形式："会話ＩＤ"-"構造ＩＤ"|格要素の対応率(構造ＩＤがユニーク)
##############################################################################
### PROCESURE 6 ##############################################################
# 処理内容：他の格構造を含まない格構造についての格要素対応率を取りだす.
# 入力形式："会話ＩＤ"-"構造ＩＤ"|格要素の対応率(構造ＩＤがユニーク)
# 出力形式："会話ＩＤ"-"構造ＩＤ"|格要素の対応率(構造ＩＤがユニーク)
# 注　　意：PROC4とPROC5の結果から、構造ＩＤが重複するものを摘出する.
#           and.cshを使用
##############################################################################
### PROCESURE 7 ##############################################################
# 処理内容：格要素対応率が1,0,その他の場合に分けて、統計をとる.
# 入力形式："会話ＩＤ"-"構造ＩＤ"|格要素の対応率(構造ＩＤがユニーク)
# 出力形式：TYPE1 件数|TYPE2 件数|TYPE3 件数
##############################################################################

# 以上.
```

(12)　　1件 各種会話情報

****
検索依頼ドキュメント
****

1．依 頼 者： 友清 研究員

2．依頼日時： 90/04/10

3．回答日時： 90/04/11

4．回 答 者： 高橋.庄山 (TIS)

5．内　　容： 会話ＩＤと会話ファイル名の対応を取る。

6．対象会話：

7．出力形式： 会話ＩＤ|ファイル名

8．検索手順： 下欄に別記する。

9．ファイル： 使用ファイルの一覧 ($SEARCH = /data/LDB/search)

個別プログラム：
$SEARCH/TOMOKIYO/TOMOKIYO1.sql # SQLスクリプト

１０．内容補足：

１１．備　　考：

１２．依頼ＩＤ：TOMOKIYO1

****
検索手順

1．エディターで次のようなファイルを作ります。

emacs script

```
lines 0
select 会話ＩＤ,ファイル名
from j_conv/
```

英語のテキスト名を出力したい場合は、３行目を次のように書き換えて下さい。

３行目　→　　from e_conv/

2．ＳＱＬコマンドで検索します。

SQL script