

“PEAR STORY” SPOKEN DATA の音声分析

Towards a prosodic analysis of spoken data:
Pauses and prosody of Pear Story narratives

石川保茂	京都外国語大学
谷村 緑	大阪外国語大学大学院
三宅弘晃	兵庫医科大学
堀内夕子	大阪キリスト教短期大学
野口ジュディー	武庫川女子大学
吉田 悦子	三重大学

Abstract

Study of a corpus of spoken data should include analysis of the prosody which can offer a wealth of information from the viewpoints of discourse analysis, psycholinguistics, pragmatics, natural language processing and language education. However, unlike written material, analysis of spoken material poses many problems, beginning with the unit of analysis itself. Fortunately, recent technological progress has led to the development of a promising speech analysis tool, multi speech (Kay, USA), which we used in this study to examine the pitch of utterances in order to try to identify units for examining spoken material. Ten native English speakers from various countries, all students at a Japanese university, were asked to view a 10-minute silent film known as the Pear Story (Chafe 1980). They were then asked to give a description about what they had viewed. Their narratives were transcribed and also subjected to analysis using multi speech to identify what could be useful as a unit of analysis. First, to identify the unit of analysis, the transcribed data by a native English speaker was classified based on semantic units, while the data by multi speech was based on identification of a pause of 200 milliseconds. Comparison of the two sets of findings showed that multi speech tended to identify about twofold more units. However, the analysis could be done automatically and quantitatively and thus was taken as the manner of identifying unit boundaries. Analysis based on this condition showed that the pitch at the start of a unit was about 1.2-fold higher than the following nucleus section, and continued to decrease to the end of the unit. The pitch level itself varied greatly with the individual. While further work is needed with more samples, our results suggest the possibility of analysis of spoken material based on pauses and prosody.

キーワード：音声分析、プロソディー、談話分析、話し言葉コーパス

“PEAR STORY” SPOKEN DATA の音声分析

1. はじめに

話し言葉コーパスは書き言葉コーパスと比べるとまだ量的には少ないが、音声分析に欠かせない重要な言語資料である。中でも、話し言葉の構造を知る手がかりとなるプロソディの研究は技術の進歩とともにより関心を集めるようになってきており、談話分析、心理言語学、語用論、自然言語処理、教育など様々な分野への応用が期待されている。しかし、話し言葉には書き言葉のような規定を当てはめることができないため、分析のための unit(発話単位)をどのように決定すべきか、その単位をどのように分類すべきか、談話分析における基本単位をどう設定すべきか、の客観的記述は難しい(前川他 2000, 堀内他 1997, 1999, Wichmann 2000, Wennerstrom 2001)。

本研究では、音声分析ツールである multi speech を利用し、分析のために収集した pear film data から unit の認定基準と、各 unit にみられる pitch(開始部分、終了部分、核の3つ周波数の値)と発言内容について考察を加える。

まず、multi speech によるデータ分析手順と、multi speech を用いて得られた音声情報に基づく客観的音声データの収集方法を説明し、次に JACET 口語英語研究会で採集した pear film data の収集方法を述べる。そして、multi speech から得られたデータ(無音区間の時間、unit 数、談話セグメント数)について記述し、unit 認定について人手で行った場合との比較を行う。さらに、各 unit の pitch より量的・質的考察を行い、最後にまとめと今後の課題を述べる。

2. 分析方法

2.1. multi speech によるデータ分析手順

この multi speech は、アメリカ合衆国の Kay 社によって開発された音声分析ツールで、波形データの多岐にわたる分析や加工を行うことを主目的としている。

本研究では、次の手順でデータを分析した。技術上の観点から、まず、音声データを 8 秒ごとに区切り、ソースとして取り込む作業を行った。また、8 秒ごとに区切ったあるデータと次のデータとの pause を測定するために、音声データ全体を取り込んだソースも作成した。次に、8 秒ごとのソースを pitch contour により分析したうえで、unit を認定した。

2.2. 転記単位

発話の音声単位は、書き言葉のように規定することは難しく、調査目的にあわせて様々な単位認定がなされている(Wichmann 2000)。その中でも特に、客観的な指標として有力な手がかりとなるのが pause である。堀内他(1997)では、「自己発話内の 400 ミリ秒以上の無音区間によって区切られた音声的連続」と定義される。また、小磯他(2000)では「原則と

して、言語音が 200 ミリ秒以上の途切れなく連続して生じている区間を転記基本単位とする。ただし、言語的な文末形式(述語の終止形や終助詞など)が存在している場合には、50 ミリ秒以上 200 ミリ秒未満の区切れであっても、その文末形式のあとで転記基本単位を分割する。」と定義される。このように定義された unit はおおむね無音の pause による切れ目におおむね相当すると考えられる。

以下で堀内他(1999)より例を見る。

(1)

00:29:168-00:31:456 G:でだいたいしたからだいたいやく

00:32:144-00:33:792 G:いっせんちぐらいのところのところ

00:33:952-00:34:144 F:はい

00:34:768-00:36:448 G:うせ<128>えとみぎにまがっ<112>てくださいⁱ

00:36:448-00:36:832 F:+あはい

00:37:024-00:37:680 G:みぎに<208>*はいⁱⁱ

00:37:520-00:38:736 F: *はい<240>はいみぎいきました

堀内他(1999)では、音声情報パワーを利用して客観的な単位としての unit(発話単位)を導入し、さらに句読点に代わる機能として、unit 内の分割をおこなう基準に各 unit 内で 100 ミリ秒以上の無音区間について< >の内側にそのミリ秒単位の時間的長さを記入する方法を採用している。本研究では、今回採集した英語データに 400 ミリ秒以上の pause 区間が極端に少なかったため、200 ミリ秒の pause を unit の切れ目と考える。

2.3. prosody と発話(単位)

前節で、pause が unit の境界認定の手がかりになることをみた。Wichmann(2000)によると、pause は、prosody の多様な音声資料の一つとされ、その他に、pitch、声の大きさ、声の質、が挙げられている。Crystal(1991)では、prosody は、pitchⁱⁱⁱ、声の大きさ、tempo, rhythm における音の変化と定義され、一方 Wennerstrom (2001)では、話し言葉の tempo, rhythm, volume, intonation を指すとされ、pitch が含まれるかどうか意見が異なる。

一般に、連続する話し言葉は、基本的な単位がいくつか結合して、談話セグメントを構成する。各談話セグメントには話題の中心(トピック)が含まれ、われわれは、それを強調

“PEAR STORY” SPOKEN DATA の音声分析

するために、声を大きくしたり、重要な語句を指す母音を長く発音したりして、他の情報と区別する、つまり、pitch pattern を変化させる(Crystal and Davy 1975)。Wichmann(2000)によると、発話の各基本単位は、前発話の構成要素と新しい発話の構成要素との統合の度合いを示す initial pitch、トピックをマークする high-pitch、談話セグメントの境界における final pitch という音声からなるとされる。言い換えると、high-pitch はトピックシフトのサインとなると考えられ、pitch の「山」と「谷」の部分の「谷」が談話セグメント境界となる。

3. pear story data 収録方法

3.1. pear story film の説明

Pear story film (Chafe 1980)は、音響効果のみの無声映画で、異なる文化背景をもつ人々がどのように同じ経験を言語化するのかを調査するために、新規にデザインされたものである。調査は 10 の異なる言語に及び、認知的な要因(e.g. 基本レベルカテゴリーの語彙選択)のほかに、テキストとコンテキスト(e.g. 誰に対して、状況、目的、スタイルのレベル)、個人差、言語特有の要因(e.g. 統語的な違いや照応の選択)などが、どのように言語化のプロセスに影響を与えるかについて、複数の研究者による調査が行われている。

3.2. 被調査者

関西の外国語大学に留学生として在籍している英語母語話者 10 名(国籍：アメリカ、ニュージーランド、カナダ、オーストラリア、イギリス)に協力をお願いした。調査は 2001 年 5 月に行われた。

3.3. データ収集の手続き

被調査者には、Pear Story のビデオを見た後、ビデオを見ていない人にその内容を伝えるつもりで語ってもらった。録音は LL 教室で行われた。

4. 分析結果

4.1. unit の認定

2.2 節で述べたように、unit 認定にはいろいろな基準が用いられているが、客観的な基準の設定は難しい。ここでは、今回採集した pear film data を用いて、人手による unit 認定と multi speech による unit 認定との比較を行う。人手による unit 認定については、関西の大学に英語教員として在籍している英語母語話者 1 名に依頼した。下の表は、unit 認定者と multi speech による unit 数を一覧にしたものである。認定者は、基本的に意味内容のまとまりがある文法的文(節)に unit を切るように課せられた。一方、multi speech では、200 ミリ秒以上の pause を unit の境界と認定した。

Data 方法	1	2	3	4	5	6	7	8	9	10	平均	標準偏 差
認定者	28	46	76	52	33	46	40	24	21	45	41.1	5.09
Multi speech	35	105	147	107	62	65	99	52	45	88	80.5	10.97

表1 認定者の直感と multi speech とによる unit の数

multi speech による unit 数は、平均で認定者のほぼ2倍となっている。標準偏差より、multi speech による unit 数の方が認定者によるものより、ばらつきが大きいことが分かる。これは、認定者は意味で unit を認定しているのに対し、multi speech では被調査者の話し方が反映するためである。以下に、具体データ(<data9>の談話)から、認定者と multi speech とで区切った unit をそれぞれ提示し、比較を行う。

(2) 認定者による unit の認定

- | | |
|--|---|
| 1. Um.. It was, based in, Latin, country, pear picking | 11. The guy in the tree was an old, middle-aged, man |
| 2. So, do you want me to describe the story? line, | 12. Obviously picking pears for his work. |
| 3. Oh I see. Um.. A guy was picking some pears | 13. Um.. the young boy, ten years old |
| 4. A young boy, riding his bike, stole the pears | 14. riding a red, bicycle, cheeky (laughter) |
| 5. um.. along the way he fell off his bike | 15. um... How would I describe the |
| 6. and... that was it (laughter) | 16. young girl, riding past him |
| 7. Three boys came and helped him pick up the pears | 17. caused him to fall off his bike |
| 8. and, there was a lot of confusion (laughter) | 18. um... and the three boy helpful boys, helped him pick up his pears |
| 9. Do you want me to describe the characters? | 19. and that's it really (laughter) |
| 10. Ah, right. Umm. Amm. | 20. I didn't see the.. there was not a lot of point (through laughter) to that. |
| | 21. Okay, thanks |

(2)の特徴としては、主に、(定・不定)動詞を中心に unit の認定がなされていることが分かる。しかし、utterance2-3 と utterance9-10 とで自問自答の形を採っているところを、utterance3

“PEAR STORY” SPOKEN DATA の音声分析

では *Oh I see* の後にさらに内容が続くと判断したのに対し、utterance10 では *Ah, right. Umm. Amm.* だけで unit として認定しており、その認定に揺れが認められることが分かる。

一方の multi speech の結果について、同じ<data9>で見てみる。

(3) multi speech による unit の認定

- | | |
|--|---|
| 1. Um.. | 24. the guy in the tree |
| 2. it was, | 25. was |
| 3. based in, | 26. an old, |
| 4. Latin, | 27. middle-aged, man |
| 5. country, | 28. obviously picking pears for his work. |
| 6. pear picking | 29. um.. |
| 7. so, do you want me to describe the
story | 30. the young boy, ten years old |
| 8. line, | 31. riding a red, |
| 9. oh I see. | 32. bicycle, |
| 10. um.. a guy was picking some pears | 33. cheeky |
| 11. a young boy, | 34. (laughter) |
| 12. riding his bike, stole the pears | 35. um... |
| 13. um.. | 36. how would I describe the |
| 14. along the way he fell off his bike | 37. young girl, |
| 15. and... | 38. riding past him caused him to fall off
his bike |
| 16. that was it (laughter) | 39. um... |
| 17. three boys came and helped him pick
up the pears | 40. and the three boy helpful boys, |
| 18. and, | 41. helped him pick up his pears |
| 19. there was a lot of confusion | 42. and that's it really |
| 20. (laughter) do you want me to describe
the characters? | 43. (laughter) I didn't see the.. |
| 21. ah, right. | 44. there was not a lot of point (through
laughter) to that. |
| 22. umm. | 45. okay, thanks. |
| 23. ammm. | |

(3)より、unit 数は認定者のものと比べるとおよそ2倍で、フィラー(*amm* など)や談話標識(*and, oh I see* など)、笑い、また、1単語(*Latin* など)も unit として認定されていることがわ

かる。さらに、1発話の語数は認定者のものが、multi speech よりも平均して多いが、語数のばらつきについては multi speech のほうが大きい。pause の置かれるところはフィラーや談話標識の前後が中心であるが、文中でも特に主語(主題)と述語(題述)間((3)の 2 と 3、24-25-26 などの例)や構文上の切れ目((3)の 11 と 12、27 と 28 など)などがあげられる。

4.2. 量的分析

4.1.では、multi speech を使用して得られたポーズの長さから、unit を認定することで、客観的データの収集が可能であることを示した。本節では、4.1.で認定した unit を基に、各 unit に見られた pitch の値から発話データの音声的記述を行う。

4.2.1. データ処理

全 10 データを下の 3 点に注目しながら、量的に処理した。

- (4) 各 unit の開始部分の周波数の平均値 O (Hz)
- (5) 各 unit の終止部分の周波数の平均値 C (Hz)
- (6) 各 unit の核(最も強調された部分)の平均値 E (Hz)

また、これらより、

- (7) 各 unit の開始部分と終止部分の差の平均値 $(C-O)/Cx100$ (%)
- (8) 各 unit の開始部分を基準とした核の強調度の平均値 $(E-O)/Ox100$ (%)

を計算した。

なお、(7)と(8)について、各 unit の開始部分の周波数とその unit 内の「高さの基準」と認識されていると仮定した。会話の参加者は unit の開始部から終止部へ向かってリアルタイムに情報を処理しているので、核部分が処理されるときには、終止部はまだ提示されていない。このような終止部を unit を貫く「基準」と捉えることは、spoken language の性質を正しく反映していると考えられない。それに対して「開始部に対して高いか低いか」という O を基準とした判断で「今現在発話されている unit の周波数」の高さを判定するというモデルは、モデルとしてより効率的なだけでなく、各 unit が終了するまで C の値を知り得ない聞き手にとってもより自然なものであり、リアルタイムの認知処理をよりうまく反映するものであると考えられる。

以上の処理により、表 2 を得た。

“PEAR STORY” SPOKEN DATA の音声分析

data	O	C	E	(C-O)/Cx100	(E-O)/Ox100
1	115.79	116.87	140.24	+0.93	+21.12
2	129.02	129.01	168.04	-0.01	+30.24
3	179.96	153.49	213.90	-14.71	+18.86
4	120.17	104.03	140.97	-13.43	+17.31
5	98.17	103.40	120.18	+5.33	+22.42
6	109.38	104.70	134.25	-4.28	+22.74
7	136.94	121.01	174.86	-11.63	+27.69
8	101.75	98.29	118.60	-3.40	+16.56
9	198.77	189.90	238.13	-4.46	+19.80
10	186.21	168.15	207.36	-9.70	+11.36

表2 量的データ処理結果

4.2.2. 結果

各 unit の開始部分と終止部分の傾き (表2 (C-O)/Ox100 を参照のこと) を取ってみると、一般的に終止部分が開始部分に比べて約5%~10%程度低い (data 3, 4, 6, 7, 8, 9, 10)。つまり、一般的に unit は「開始部分が高く、終わりに向かって低くなる」という右肩下がりの傾向を持っていることがわかる。しかしながら、data 1 や data 2 のように開始~最終部分の高さがあまり変わらない話し方や、data 5 のように終止部分の方が逆に高くなるような話し方も少数ではあるが見られ、必ずしも右肩下がりであるとは限らないことがわかる。

各 unit の nucleus (核) (表2 (E-O)/Ox100 を参照のこと) となる部分は、開始部のおよそ1.2倍(+20%)の高さを示し、この値は話者 (data の違い) に大きく左右されることはない。核としたい部分を核とするためには、ある部分の音節を「目立たせる」という目的を果たすために必要十分な周波数の差を得ることができればよく、それがおよそ+20%という値であると考えられる。

4.3. 質的分析

4.3.1. データ処理

4.2.の量的研究で最も平均的な値を出していた<data4>を扱う。ただし、開始部分データの信頼性が明らかに怪しい unit 29(O₂₉=11.15)は除外した。

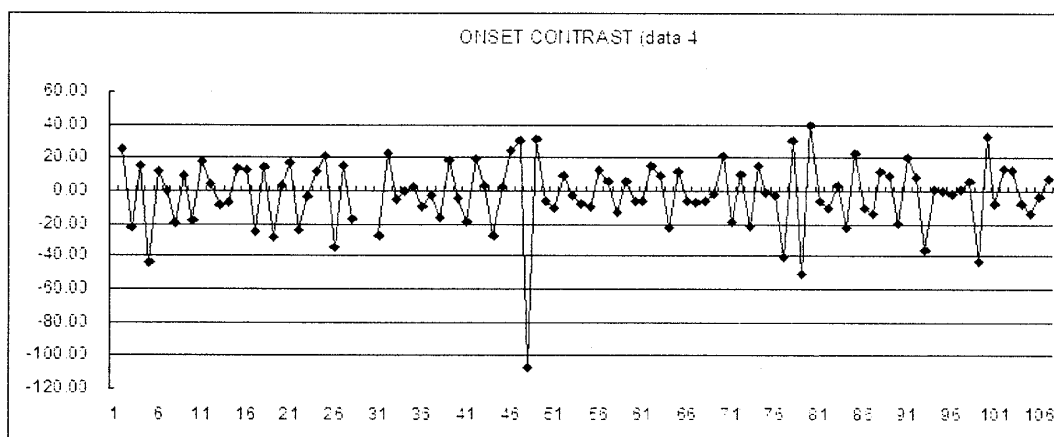
このデータを下の3点に注目しながら処理した。

JACET関西紀要 第7号

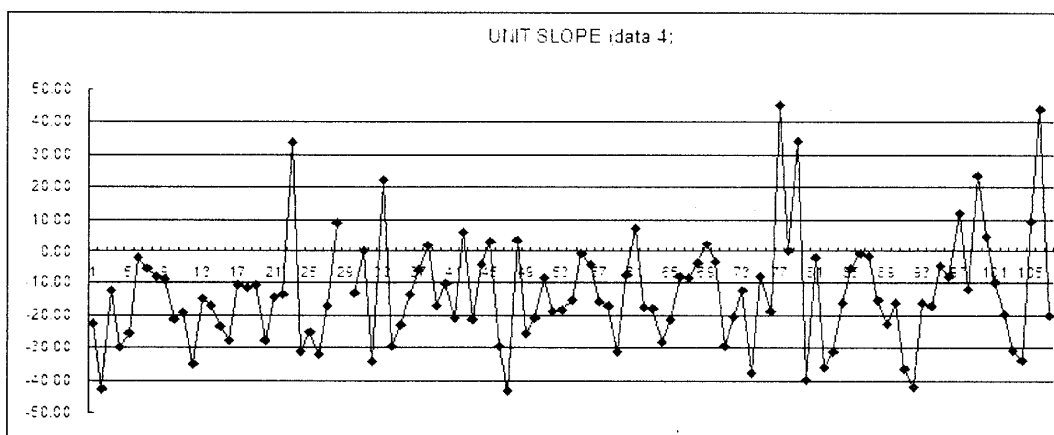
- (9) unit n の開始部分の周波数と一つ前のそれとの差分 $(O_n - O_{n-1}) / O_{n-1} \times 100$ (%)
- (10) unit n の開始部分と終止部分の差分 $(C_n - O) / O_n \times 100$ (%)
- (11) unit n の開始部分を基準とした核の強調度 $(E_n - O_n) / O_n \times 100$ (%)

を計算した。(9)の値を取った理由は、各 unit の基準周波数となる値 O の連続的な変化を調べるためである。(なお、(10)と(11)を求める意義については、4.2.の量的研究で行った考察に準ずるので、そこを参照されたい。)

以上の処理により、グラフ 1~3 を得た。

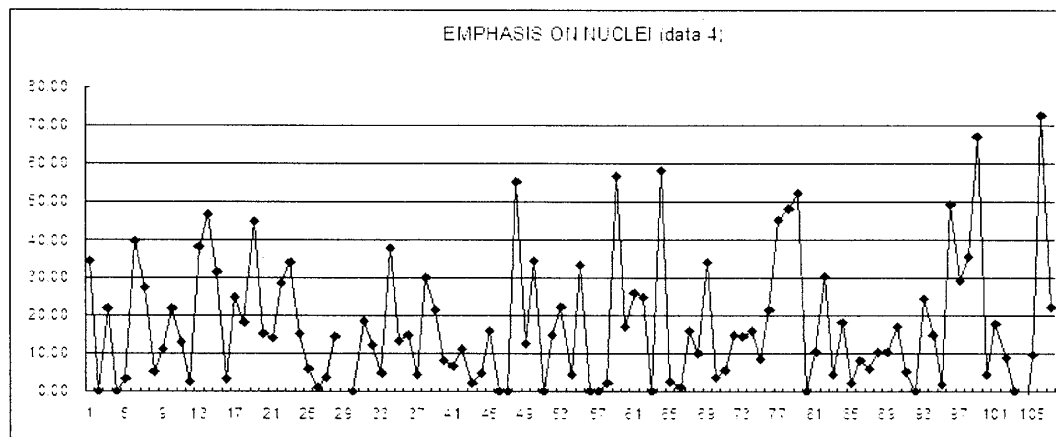


グラフ 1 各 unit の開始部分の周波数と一つ前のそれとの差分



グラフ 2 各 unit の開始部分と終止部分の差分

“PEAR STORY” SPOKEN DATA の音声分析



グラフ 3 各 unit の開始部分を基準とした核の強調度

4.3.2. 結果

上の質的分析の結果、最も興味深いものは、グラフ1である。各 unit の開始部分の周波数をプロットすると、データ全体を通じて右肩下りの傾向が見えるが、この「一般的な傾向」(general tendency of downfall in intonational contour)に反する部分に注目すると、その unit (とその中に含まれる情報)が discourse の中で重要な意味を持っていることがわかる。unit 31-44 を例にとる。

unit	CONTRAST	CONTENTS
31	-26.74	he begins to
32	+22.93	uh take one pear
33	-4.77	then he looks again
34	0	and decides to take a whole basket of pears
35	+2.28	the boy places the uh
36	-9.3	uh basket of pears
37	-2.13	onto the rack
38	-16.44	uh on the front of his bicycle
39	+18.59	the man in the tree
40	-4.39	uh does not notice the boy at all

41	-18.94	while he rides away
42	+19.46	the boy makes his way down the road
43	+3.3	trying to balance the basket onto his on the rack
44	-27.27	uh

表3 各 unit と発言内容対照表 (太字部は筆者による)

表3から、unit 32, 39, 42の開始部分が前unit比およそ1.2倍(+20%)と際立っていることがわかる。unit 32ではone pear、unit 39ではthe man in the tree、unit 42では(to) make his way down the roadという新たな情報が談話の中に導入され、それ以降これらの新たな情報が談話の主題(話題の中心)となっていることに注目したい。units 32-37を見ると、unit 32で与えられたpear(s)がunit 34, 36でも繰り返されている。しかし、開始部分が右肩下がりという一般的傾向に明らかに反して引き上げられているunit 39以降では、basketで、前のunitとの結束性(cohesion)は保持されているものの、pear(s)の使用は全く見られず、新たに提示されたthe man (とthe boyとの関係性)へと主題が移っていることがわかる。このことは、これら複数の音声上の単位が、談話におけるより大きな主題を表す単位を構成していることの傍証と考えられる。

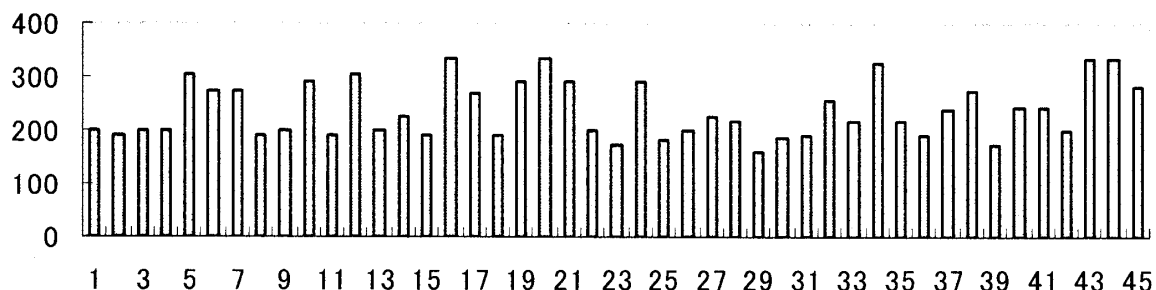
つまり、この音声単位の開始部分における高周波数は、音声単位(tone unit)をより大きな内容(話題)の単位(larger unit)にまとめあげるものであると言えるであろう。あるいは、逆に「談話における主題が、高い周波数を持つ音声単位の開始によって示されている」とも言ってもよい。ただし、unit31-32のように、to不定詞のtoと不定形takeの間に境界が割り込むなど、内容(話題)上の単位と文法上の単位が一致するとは限らない。

5. 今後の課題

本節では、より大きな内容の単位(談話セグメント)の境界、つまり、核の強調度の周波数で低いところ、「谷」が境界として認定できる根拠として、どのような言語的特徴を見られるかを指摘し、今後の課題としたい。

“PEAR STORY” SPOKEN DATA の音声分析

NUCLEI (data9)



グラフ 4 各 unit の核と談話セグメント

グラフ 4 より周波数の値からトピックの境界を認定すると、unit 8, 11, 13, 15, 18, 23, 25, 29, 33, 36, 39, 42, 45 がそれに当たる。(3)の<data9>から一部を取り上げる。

(3)'

- | | |
|--|--|
| 18. <u>and</u> , (接続詞) | 31. riding a red, |
| 19. there was a lot of confusion | 32. bicycle, |
| 20. (laughter) do you want me to describe
the characters? | 33. <u>cheeky</u> (評価) |
| 21. ah, right. | 34. (laughter) |
| 22. umm. | 35. um... |
| 23. <u>amm.</u> (言いよどみ) | 36. <u>how would I describe the</u> (自問) |
| 24. the guy in the tree | 37. young girl, |
| 25. <u>was</u> (トピックの同定) | 38. riding past him caused him to fall off
his bike |
| 26. an old, | 39. <u>um...</u> (言いよどみ) |
| 27. middle-aged, man | 40. and the three boy helpful boys, |
| 28. obviously picking pears for his work. | 41. helped him pick up his pears |
| 29. <u>um..</u> (言いよどみ) | 42. <u>and that's it really</u> (コーダ) |
| 30. the young boy, ten years old | |

(3)'より、談話セグメント境界にみられる言語的特徴として、間をつなぐ接続詞や言いよどみが挙げられる。また、ストーリーラインと直接関係しない、話者の問いかけ、評価など

も境界に見られる。このように音声情報から得られた談話セグメント境界にはある特定の言語表現出現することが予想される。同様の分析を全発話者に対して行った。表4はその結果である。

機能	頻度	機能	頻度	機能	頻度
間つなぎ (接続詞)	38	トピックの導入	11	コーダ (終わりの合図)	2
間つなぎ (いいよどみ)	15	補足追加	6	自問	2
並列	28	対比	5	時間設定	1
トピックの同定	23	評価	5	描写	1
場所設定 (前置詞句、従属節)	18	言い換え	4	その他	3
修飾 (形容詞句、関係節)	15	話者の感想	3		
説明 (題述)	14	繰り返し (解説)	2		

表4 全発話者の談話セグメント境界に見られる言語的特徴

一番頻度の高い談話セグメント境界の言語特徴には、間をつなぐ接続詞と言いよどみが挙げられる。これら二つをあわせると、全体の約27%にあたり、4回に1回の割合で談話セグメントの境界を記す役割を果たしていることがわかる。その他としては、並列、トピックの同定、場所設定、修飾、説明(題述)などが挙げられる。これらの例に共通する点は、下線部と前発話とは統合されているが、次の発話では新たな発話単位が始まるという点で、前後の発話間をつなぐ単独の接続詞や言いよどみとは異なるものである。

6. 結論

本研究では、multi speech を用いたデータ処理に基づき、量的・質的分析を行った。200ミリ秒以上のポーズを unit の境界と認定することで、音声情報のみに基づく客観的データの提示が可能であることを示した。量的分析として、各 unit に見られる pitch の周波数値計測により、(1)一般的には unit は「開始部分が高く、終わりに向かって低くなる」という右肩下がりの傾向があるがその傾きは話者に大きく左右されること、(2)各 unit の nucleus

“PEAR STORY” SPOKEN DATA の音声分析

となる部分は開始部分のおよそ 1.2 倍の高さを示し、この値は話者に大きく左右されないことが指摘された。質的分析としては、談話における主題が、前後の unit の開始部分よりも高い周波数を持つ unit の開始によって示されることが判明した。その差異はおよそ 1.2 倍であった。なお、本研究で収集されたデータのなかには、音声以外のノイズの周波数が記録されている可能性が捨てきれない。こういった「異常な」データを排除するには現段階では人手作業が不可欠であるが、コーパスを構築する際には統計的な裏付けをもとにして自動的に異常を検出できるよう、エラー検出用の対照データとして保存・調査を進める必要がある。また、音声情報による談話セグメント境界と言語特徴との関係について更に研究を進める必要もある。このような課題を克服しコーパス化が実現した際には、構築したコーパスの英語教育への様々な活用法について検討を進めて行きたい。

参考文献

Chafe, W. L., (ed.) (1980). *The pear stories*. Norwood: N. J. Ablex

Crystal, D. (1991). *A dictionary of linguistics and phonetics* (3rd ed) Blackwell.

Crystal, D. and D. Davy. (1975). *Advanced conversational English*. Longman.

小磯花絵・土屋菜穂子・間淵洋子・斉藤美紀・籠宮隆之・菊池英明・前川喜久雄(2000)「日本語話し言葉コーパス」における書き起こしの方法とその基準について『日本語科学』9, 43-58.

堀内靖雄, 吉野文, 仲真紀子, 土屋俊, 市川熹. (1997). 「千葉大学地図課題対話コーパスプロジェクト」『千葉大学工学部研究報告』 Vol. 48. No2. pp33-60.

堀内靖雄・中野有紀子・小磯花絵・石崎雅人・鈴木浩之・岡田美智男・仲真紀子・土屋俊・市川あきら. (1999). 「日本語地図課題対話コーパスの設計と特徴」『人工知能学会誌』14(2), 261-272.

前川喜久雄・籠宮隆之・小磯花絵・小椋秀樹・菊池英明. (2000). 「日本語話し言葉コーパスの設計」『音声研究』4(2), 51-61.

Wichmann, A. (2000). *Intonation in text and discourse*. Essex: Longman.

Wennerstrom, A. (2001). *The music of everyday speech: prosody and discourse. analysis*. Oxford University Press.

ⁱ +は、ある話者の発話単位の終了直前に相手話者が発話単位を開始する場合、挿入される。

ⁱⁱ *は、重複の開始場所に挿入される。また、重複の場合、空行は設けない。

ⁱⁱⁱ patterns of movement(falls, rises, fall-rises, rise-falls)などを指す。tone と同意。