

文節構造と韻律情報の関係についての考察

7N-7

大深悦子

日本アイ・ビー・エム株式会社

金子 宏

サイエンス・インスティテュート

1. はじめに

テキスト音声合成における話調成分(ポーズ長, 話調の立て直し度)の設定には, 一般に,

- (1) 一息に発声する長さ
- (2) 文節間の修飾タイプ(連用修飾or連体修飾)
- (3) 文節構造(構文情報)

などが関連すると言われている。本稿では, このうち(3)文節構造に重点を置いて, これらと話調成分情報の関連を実音声の分析により調べたので, ここに報告する。

2. 構文情報を表わす指標

韻律制御のための文節構造の指標としては, 様々な方式が提案されている。[1]-[4]

人間が文章を声に出して読む場合, これから読もうとする文節が

- (a) 新しい意味のまとまりとして始まるのか否か,
- (b) 次の文節に直接係るのか否か,

などを手掛りに, 抑揚をつけるのではないかと考えられる。ここでは, これらの要因を含む指標として, 係り受け関係にある文節を, かつこ対で囲んだ際, 文節境界に現われる“かつこの数とパターン”を取り上げた。

例: ((かわいい 花子が)(元気に 登校した))

かつこ数: 0 2 0
パターン: -) (-

3. 実験

3.1 音声試料

男性話者1名が叙述文100文を約6モーラ/秒で発声したものを音声試料とした。

3.2 韻律レベル

韻律制御のために, 上記の試料中の文節境界の話調を次の5つのレベルに分類した。

レベル	記号	話調立直し	ポーズ長	
1	-	無	0	一息に読む
2	F	有	0	話調立直し
3	(S)P	有	< 100ms	小ポーズ挿入
4	(M)P	有	100ms ~ 300ms	中ポーズ挿入
5	(L)P	有	>= 300ms	大ポーズ挿入

4. 結果と考察

結果を表1-3にまとめる。なお, ポーズ長は発声速度6モーラ/秒に規格化している。これから以下の点が明らかになった。

- (1) かつこ数0の切れ目は続けて発声される。

ただし, 発声モーラ数が一定値(3-3.5秒で話せる程度)を超える場合は, ポーズが挿入される。

<かつこ数0でポーズが入った例>

((私たちの 音声合成システムの)できあが~)
7モーラ ↑ 13モーラ

- (2) かつこ数3以上の切れ目はポーズが挿入される。

5人の被験者に叙述文57文を与えてよんでもらったところ, かつこ数3以上の切れ目(31か所)のうち, 5人ともポーズ挿入(11か所:35%), 4人(12か所:38%), 3人(5か所:16%)となった。5人の発声者のうち半数以上が対象箇所の90%以上にポーズを入れていることになり, 個人差を考慮しても, かつこ数3以上の切れ目は切って発声するとみてよいと思われる。

<かつこ数3以上でポーズの入らなかつた例>

((私が(あなたに(あれを(お土産に求めた))))店)の住所を)置きます)。 ↑

理由として, “求めた”が連体修飾で結合がつよいこと, 終止形と同形なのでここで切ると文が終わったように感じられること, “店の”が3モーラなので続けることが可能なこと, などが考えられる。これより, 修飾タイプによっては, 係り受け構造を超えて影響力を持つものがあることがわかる。

- (3) かつこ数1, かつこ数2の切れ目の取扱い

(i) かつこパターンによる切れ方の差

切れ方の強い順に並べると, 表2のようになる。これより,
)(: 大ポーズ
((: 中~大ポーズ

-) : 中ポーズ
- (: 話調立て直し
-) : 続けて発声～話調立て直し

に、それぞれ対応していることがわかる。これは、左かっこが新しい意味のまとまりが始まるのを表わすのに対して、右かっこは、意味のまとまりを閉じて次の文節にかけることを表わしているため、切れ方が弱いものと思われる。また「(」の切れ方の強いのは、これが、意味のまとまりを閉じるのと開くという2つの役割を担っていること、「(」に比べて一般に発声モーラ数の大きいことなどが考えられる。

(II) 修飾タイプによる影響

表3をみると、修飾タイプによって、同じかっこタイプのもので切れ方にばらつきがあるのがわかる。属する韻律レベルが平均と著しく異なっているものは、以下のとおりである。

[かっこ数1]

- 3a. 用言連体形 (平均2.5) - 0.8
- 1. 接続詞 + 0.5
- 4f. 強調の副助詞 + 1.1
- 2. 並列'と' + 1.5
- 6. 接続助詞 + 2.3

[かっこ数2]

- 5. 副詞 (平均3.7) - 1.3

5. おわりに

文節構造の指標として、かっこ数とパターンを使用することを提案し、実音声の分析によって、この指標、修飾タイプと、話調成分情報との関連を調べた。これをもとに、話調成分付与規則を作り、いくつかの文に適用して実際に聞いてみたところ、かなり自然な抑揚が得られた。今後は、この規則の改良を進めるとともに、今回除外した疑問文、命令文の抑揚や、話調成分に与える発声速度の影響などを検討する予定である。

参考文献

[1]中島, 壁谷「文節構造に着目した文音調パターンに関する一考察」音学会講論集 2-2-8 (1984-03)
 [2]箱田, 佐藤「文音声合成における音調規則」信学論誌 Vol. J63-DNo.9 (1980-09)
 [3]藤崎, 広瀬他「統語構造を利用した韻律記号導出アルゴリズム」音学会講論集 2-2-12 (1986-03)
 [4]鍋木他「日本語音声の規則合成システム-カナ文字から音声への変換」音学会音声研資料 S85-44 (1985)

表1 かっこ数, 修飾タイプの話調成分への影響 (件数)

かっこ数 韻律レベル	0		1		2		3		4以上	
	-	F P	-	F P	-	F P	-	F P	-	F P
修飾タイプ										
1 接続詞	1	1	3	2	5					
2 並列'と'	1	1	5		5					
3a 連体形	18	1	9	4	3	1	1	4	1	1
b 'の'	39	10	2	5	4	3	6	1		
4a 'が'	3		12	2	7	2	1	7	3	6
b 'の'	5		2	1	2	5				
c 'を'	7	2	8	6	6	1	5	1		3
d 'に'	15	1	1	6	4	5	2	1	3	2
e 'と'	2	1	1	3	1	5		1		
f 強調副助	2		2	3	1	4				
g 'で'	1		1	4	4	2	4	1		1
h その他			1	1	1	1				
5 副詞	33	4	1	2	4	3	1	5	3	
6 接続助詞				5	1	6		8		
7 提示'は'	2	1	3	4	6	1	14	3		4
8 読点			8		2		5			2

-: 続けて発声, F: 話調立て直し, P: ポーズ挿入

表2 かっこパターンの韻律レベルへの影響 (件数)

	1	2	3	4	5	計	平均韻律レベル	σ
) (0	4	3	12	18	37	4.2	1.00
((3	7	5	10	12	37	3.6	1.34
))	4	2	3	14	0	23	3.2	1.19
							平均	3.7
(18	26	5	16	9	74	2.6	1.38
)	33	15	5	22	4	79	2.4	1.40
							平均	2.5
計	58	54	21	74	43	250	全平均	2.9

表3 修飾タイプと韻律レベルの関係

修飾タイプ	かっこ数1		かっこ数2	
	平均韻律レベル	数	平均韻律レベル	数
1 接続詞	3.0	5	3.6	5
2 並列'と'	4.0	5	4.6	5
3a 連体形	1.7	16	3.5	6
b 'の'(連体修飾)	2.3	12	4.7	6
4a 'が'(連用修飾)	2.0	21	3.4	10
b 'の' (")	2.2	5	4.4	5
c 'を' (")	2.2	20	3.7	6
d 'に' (")	2.1	15	2.7	6
e 'と' (")	2.2	5	4.4	5
f 強調副助(")	3.6	5	3.8	5
g 'で' (")	2.8	9	3.5	6
5 副詞	2.6	9	2.4	9
6 接続助詞	4.8	5	3.6	7
7 提示'は'	2.8	13	4.3	15