

An empirical study on the way to create
effective audio-visual information using telops
and sound effects

金, 基弘

<https://doi.org/10.15017/459597>

出版情報 : 九州大学, 2006, 博士 (芸術工学), 課程博士
バージョン :
権利関係 :

第 2 章 テロップと効果音の時間構造と 変化パターンの類似に基づく構 造的調和

2.1 はじめに

第 1 章で述べたように、これまでに行われてきた音と映像の構造的調和に関する研究は、そのほとんどが時間構造としての視覚事象と聴覚事象のアクセントの同期の影響を対象としたものである (e. g., Iwamiya, Sugano, & Kouda, 2000; Lipscomb 1996, 1999; Sugano & Iwamiya, 1998)。しかし、音と映像の構造的調和は、視覚事象と聴覚事象の動きや変化パターンの類似の間にも成立しうるものである (e. g., Eitan & Granot, 2004, 2006; Iwamiya & Ozaki, 2004; Lipscomb & Kim, 2004)。それにもかかわらず、音と映像の変化パターンの類似による構造的調和の効果はあまり研究対象とされることはなかった。

各種の表現様式を必要とする映像メディアでは、各種の構造的調和の効果を必要とする。とりわけ、短時間に映像内容を印象づけるため、テロップ・プレゼンテーションにおいては、その必要度は高い。テロップと効果音の時間構造の類似のみならず変化パターンの類似も、音と映像の構造的調和を生み出すのに効果的であると考えられる。本章では、各種の構造的側面を構成するテロップと効果音の組み合わせを用い、二つの側面の音と映像の構造的調和の効果を明らかにすることを目的とする。

2.2 実験 1：テレビ番組のテロップと効果音を用いた実験

実験の第 1 段階として、制作者が意図したテロップと効果音を実験刺激として用い、音と映像の主観的調和の要因を探ることを目的に、SD 法による印象評定実験を行った。

2.2.1 実験方法

実験素材は、テロップがよく使われている日本のテレビ番組を VHS 方式のビデオテープに録画/録音し、多用されている典型的なパターンの中からテロップと効果音が同時に提示された部分（以下「オリジナル条件」と記す）を 8 ヶ所選んだ。抜粋した番組を表 2.1 に示す。

実験素材は、ビデオプレーヤーからコンピュータに取り込んでデジタル変換した。映像は Adobe 社の Premiere 6.0 で、音はデジオン社の DigionSound 2 で編集を行った。意図していない要因が実験の結果に及ぼす影響を避けるため、できる限りテロップだけが映っている部分を選び、背景音楽と音声の入った部分は編集で取り除いた。そして、映像の場合は 1 秒あたり 30 frame のフル・フレームに、音の場合はサンプリング周波数を 44.1 kHz、量子化ビットを 16 bit、stereo 形式の WAVE ファイルに書き出した。実験に用いた 8 種類のテロップと効果音の特徴を、それぞれ図 2.1 と表 2.2 に示す。

印象評定実験には、これらの 8 種類のテロップと 8 種類の効果音を組み合わせた 64 種類（8 種類のオリジナル条件を含む）の視聴覚刺激を用いた。

表 2.1 実験 1 の実験素材として用いたテレビ番組

映像刺激	抜粋番組	放送局	放映日(曜日)
Telop 1	特命リサーチ 200X II	FBS(福岡放送)	2003 年 2 月 16 日(日)
Telop 2	特命リサーチ 200X II	FBS(福岡放送)	2003 年 2 月 16 日(日)
Telop 3	特命リサーチ 200X II	FBS(福岡放送)	2003 年 2 月 16 日(日)
Telop 4	特命リサーチ 200X II	FBS(福岡放送)	2003 年 2 月 16 日(日)
Telop 5	特命リサーチ 200X II	FBS(福岡放送)	2003 年 2 月 16 日(日)
Telop 6	イマジン!	KBC(九州朝日放送)	2003 年 2 月 11 日(火)
Telop 7	イマジン!	KBC(九州朝日放送)	2003 年 2 月 11 日(火)
Telop 8	ロンドンハーツ	KBC(九州朝日放送)	2003 年 2 月 11 日(火)

表 2.2 実験 1 の音刺激として用いた各効果音の持続時間(mm), 呈示音圧レベル(dB(A))
とその特徴

音刺激	持続時間	呈示音圧レベル	特徴
Sound 1	2433	71	素早く鳴る 17 個の金属性の断続音
Sound 2	2233	63	速く立ち上がってゆっくりと減衰してゆくノイズ
Sound 3	266	62	500 ms の無音の後の最も短い破裂音(noise burst)
Sound 4	1000	63	滑らかな立ち上がり立ち下がりを持つ破裂音
Sound 5	733	61	速く立ち上がるタムタムのような単音
Sound 6	2300	62	ウインドチャイムのような短音の流れ
Sound 7	1666	66	自然的短音階の音高が上昇するメロディ
Sound 8	1933	58	2 連打の和太鼓の音



(0 ms).....(99 ms).....(231 ms).....(2264 ms).....(2433 ms)

a. Telop 1: 画面の下の部分に17個の文字が等時間間隔で提示される ("Typewriter")



(0 ms).....(594 ms).....(858 ms).....(1396 ms).....(2233 ms)

b. Telop 2: 画面の中央で一回ゆっくりと光ってから消える (Fade in & out)



(33 ms).....(594 ms).....(924 ms).....(1594 ms).....(3933 ms)

c. Telop 3: 右側の画面外から影を落としながら滑り込んできて左側に停止する (Slide in)



(66 ms).....(231 ms).....(660 ms).....(1132 ms).....(3166 ms)

d. Telop 4: 三つの変化パターンが組み合わされている (Fly in + Fade in & out + Zoom in)

CONTINUED ON NEXT PAGE



(0 ms)……………(66 ms)………(2924 ms)

e. Telop 5: 一回提示されてから消えるまでまったく変化しない (Still)



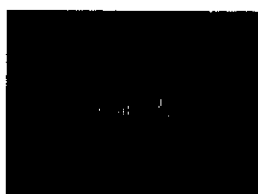
(0 ms)……………(2933 ms)

f. Telop 6: 画面の中央で前の方へ徐々に拡大してくる (Zoom in)



(66 ms)……………(792 ms)……………(1132 ms)……………(1462 ms)………(3900 ms)

g. Telop 7: 画面の中央で徐々に明るくなる (Fade in)



(0 ms)………(2166 ms)

h. Telop 8: 一回提示されてから消えるまでまったく変化しない (Still)

被験者の課題は、表 2.3 の 10 対の両極の印象評定尺度を用い、64 種類の視聴覚刺激の印象を評定することである。視聴覚刺激の主観的調和に関する評定には、「調和していないー調和している」の尺度を用いて行った。これらの印象評定尺度は、音と映像の相互作用に関する先行研究を参考に、本実験に用いる刺激の特徴を考慮したうえで選んだ。なお、実験に用いたすべての印象評定尺度は、7 段階に構成し（図 2.2-b）、「どちらでもない」を 3 に、両極を 0 と 6 に点数化した。

実験は、防音室内の照明を消した状態で行った。実験室内の暗騒音は 36 dB (A 特性音圧レベル) であった。照度は、照度計 (MINOLTA T1) で測定した結果、0.51 lx であった。被験者の眼球からディスプレイまでの距離は約 70 cm であった。図 2.2-a に、実験の様子を示す。

刺激の呈示と印象評定は、パーソナルコンピュータ (IBM Think Centre A30P) を用いて行った。映像刺激は、17 inch のディスプレイ (IIYAMA MT-8617E) の画面上に呈示した。音刺激は、オーディオ・インターフェイス (SOUND BLASTER AuDiGY2) とヘッドフォン (Sony MDR-CD7) を介して、被験者の両耳に呈示した。

表 2.2 に示した各効果音の呈示音圧レベル (等価騒音レベル) は、実験前に被験者ごとに音刺激を聴くのに最適であると思われるレベルに調整させた後、人工耳 (Brüel & Kjaer Type 4153) と普通騒音計 (Rion NA-29) を用いて測定して平均した値である。画面のコントラストも、実験前に被験者に見やすいレベルに調整させた。

すべての刺激と 7 段階の印象評定尺度は、ディスプレイの画面上にランダムな順で呈示し、マウス装置を用いて評定を行わせた。被験者にはすべての視聴覚刺激に対し、1 回呈示が終わってから評定するように教示した。また、評定の難しかった場合は何度も繰り返して視聴できるようにした。被験者には実験環境と課題に慣れさせるため、本試行に入る前に 2 回の練習試行を行わせた。なお、被験者の負担を減らすため、本試行の途中に 1 回の休憩を挟んだ。実験後は、視聴覚刺激の調和判断に関する被験者の意識を自由記述の形式で回答を求めた。

被験者は、日本語を母語とする 13 名 (男性 11 名, 女性 2 名) の九州芸術工科大学の音響設計学科に在学する大学生および大学院生である。年齢は 21 歳から 29 歳 (平均年齢 23.5 歳) である。すべての被験者は正常な視力 (矯正視を含む) を有しており、映像に関する特別な訓練を受けていない。なお、被験者たちは本実験に用いた刺激を事前

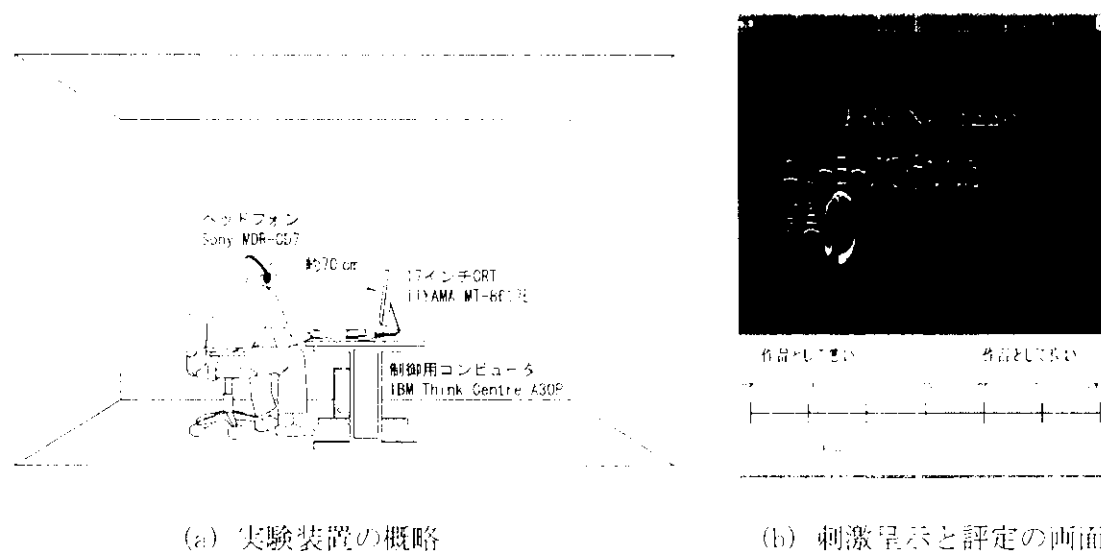


図 2.2 実験装置の概略と印象評定実験の様子

に視聴していなかった。視聴覚刺激の調程に関する実験には 13 名が参加しており、そのうちの 9 名が視聴覚刺激の印象に関する実験に参加した。

2.2.2 結果と考察

視聴覚刺激の印象評定実験より得られたデータをもとに平均評定値を求め、評定尺度を変量とした主成分分析を行った。バリマックス回転後の主成分係数行列を表 2.3 に示す。

第 1 主成分は、「悪いー良い」「つまらないー面白い」「分かり難いー分かり易い」などの尺度に負荷が高く「評価力動」主成分と解釈した。第 1 主成分には「インパクトのないーインパクトのある」「弱々しいー力強い」などの尺度も負荷が高く、力動性を含んだ主成分である。第 2 主成分は、「まじふれたーユニークな」「単調なー変化や変動のある」などの尺度に負荷が高く、「活動性」主成分と解釈した。また、これらの 2 主成分に対する各視聴覚刺激の主成分得点を求め、図 2.3 には「評価力動」主成分得点を、図 2.4 には「活動性」主成分得点を示す。

表 2.3 実験 1 の視聴覚刺激の印象評定実験より得られた主成分係数行列

印象評定尺度	主成分	
	評価/力動	活動
悪いー良い	.933	-.117
つまらないー面白い	.900	.117
分かり難いー分かり易い	.893	-.285
ゆったりとしたースピード感のある	.865	.385
インパクトのないーインパクトのある	.865	.385
暗いー明るい	.694	.009
弱々しいー力強い	.691	.303
ありふれたーユニークな	-.224	.800
鈍いー鋭い	.205	.734
単調なー変動や変化のある	.179	.664
寄与率 (%)	50.6	21.3

図 2.4 の「活動」主成分得点の場合, Telop 3, Telop 5, Telop 6 のように, 全般的にテロップに各効果音を組み合わせても活動感の変化が極端に変わることはない。つまり, 視聴覚刺激の活動感は音刺激より映像刺激によって印象付けられる傾向がある。一方, 図 2.3 の「評価/力動」主成分得点の場合, 図 2.4 とは違い, 各テロップに効果音を組み合わせることによって評価/力動感が大きく変わる傾向がある。また, Telop 1 と Sound 1 のように, 特定のテロップと効果音が組み合わせられることによって評価/力動感が得られたり, Telop 7 のように映像刺激によって印象づけられたりする傾向もある。

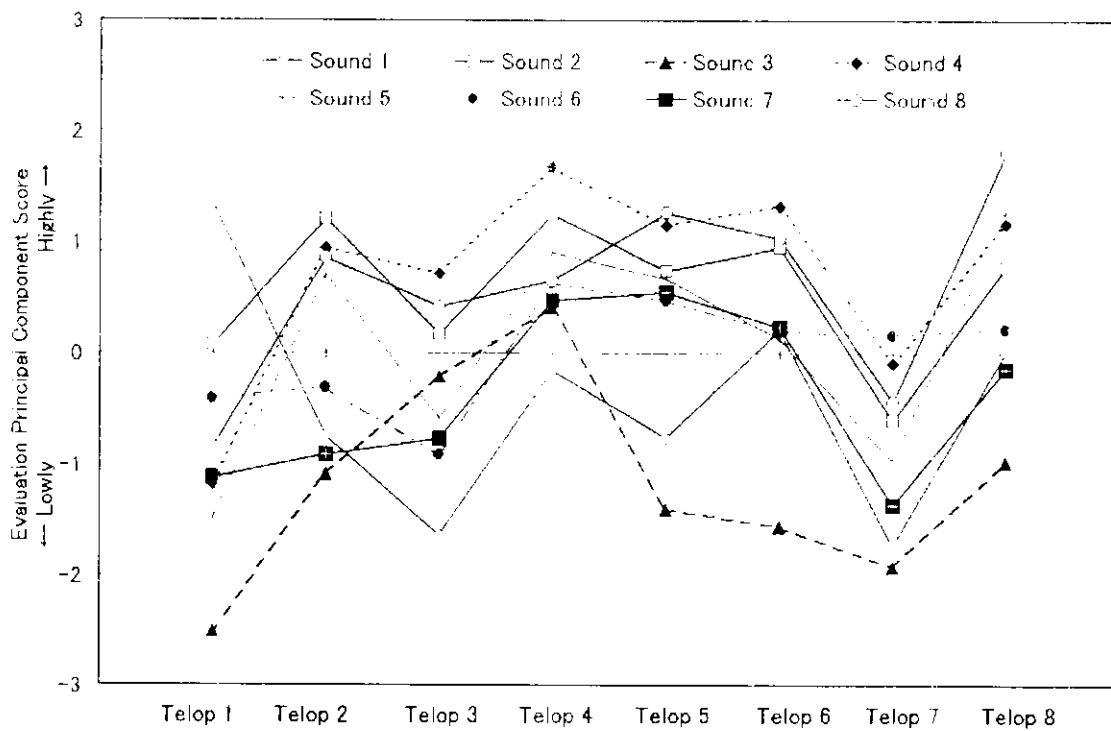


図 2.3 実験 1 の各視聴覚刺激の評価/力動主成分得点

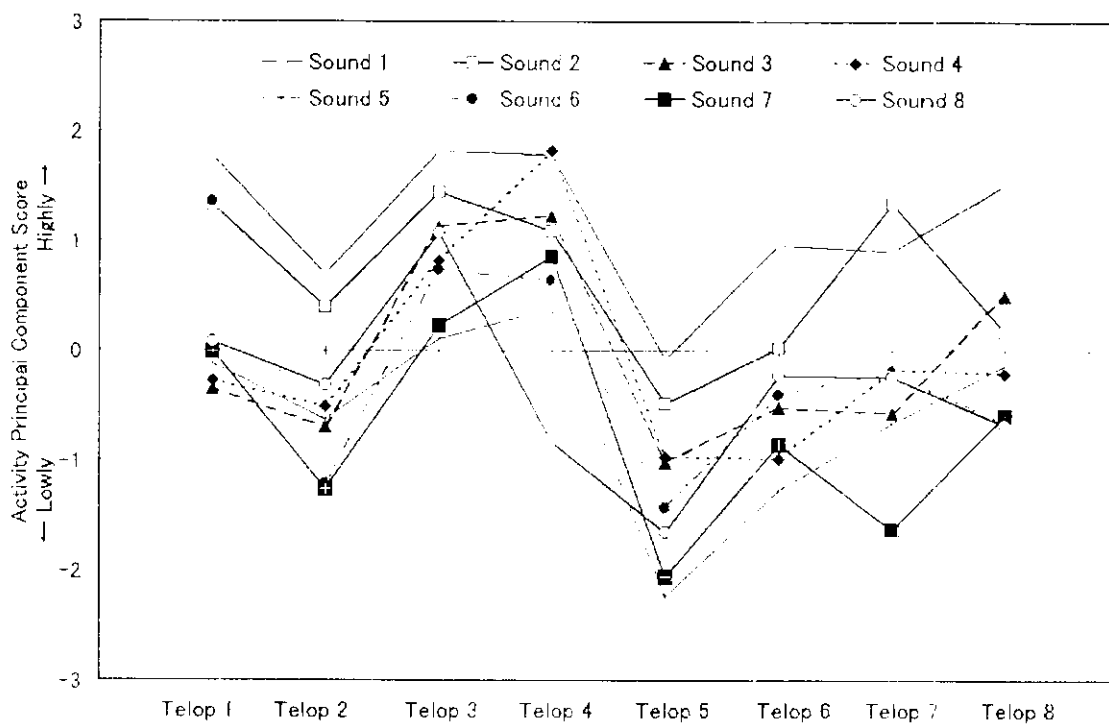


図 2.4 実験 1 の各視聴覚刺激の活動主成分得点

各視聴覚刺激の調和尺度に対する評定値を被験者間で平均した平均調和評定値を求めた。実験要因（映像 8 水準×音 8 水準）の分析のために反復測定による二元配置の分散分析を行った結果、映像刺激の主効果 [$F_{(7,84)} = 9.61, p < .01$], 音刺激の主効果 [$F_{(7,84)} = 21.09, p < .01$], 映像刺激×音刺激の交互作用 [$F_{(49,568)} = 5.89, p < .01$] が、それぞれ統計的に有意であった。二つの要因間に交互作用が存在しており、各要因の水準間の差（主効果）の検定はあまり意味を持たないため、水準ごとに一元配置の分散分析を行った（石村, 1992, 2002）。その結果, Telop 1 [$F_{(7,84)} = 16.39, p < .01$], Telop 2 [$F_{(7,84)} = 3.96, p < .01$], Telop 3 [$F_{(7,84)} = 3.94, p < .01$], Telop 4 [$F_{(7,84)} = 5.93, p < .01$], Telop 5 [$F_{(7,84)} = 14.35, p < .01$], Telop 6 [$F_{(7,84)} = 11.06, p < .01$], Telop 7 [$F_{(7,84)} = 11.66, p < .01$], Telop 8 [$F_{(7,84)} = 12.69, p < .01$] のように、各映像刺激における効果音の主効果がすべて統計的に有意であった。図 2.5 に、映像刺激ごとに各音刺激との組み合わせに対する平均調和評定値を示す。

図 2.5 より、各映像刺激においてオリジナル条件の方が意図的に調和を崩すために別のテロップの効果音と組み合わせた条件（以下「組み合わせ条件」と記す）より調和度が高い傾向がある。便宜的に、平均調和評定値が 4 以上の視聴覚刺激を調和度の高いものと定義すると、オリジナル条件の Telop 1 と Sound 1, Telop 4 と Sound 4, Telop 6 と Sound 6, Telop 7 と Sound 7, Telop 8 と Sound 8 の組み合わせは、各映像刺激において他の組み合わせ条件より最も調和度が高い。Telop 5 と Telop 2 の場合、オリジナル条件の調和度はかなり高いが、オリジナル条件よりも調和度の高い組み合わせ条件もある。Telop 3 では、オリジナル条件も組み合わせ条件も調和度が低い。

Telop 1 と Sound 1 のオリジナル条件は、全視聴覚刺激の中で最も調和度の高い組み合わせである。「タイプライター」のように、132 ms 間隔で規則的に 1 文字ずつ素早く提示されるアクセント感のあるテロップに時間的に同期したアクセント感のある効果音が組み合わせられ、非常に調和していると評定されたと考えられる。一方、Telop 1 に対する組み合わせ条件ではすべて調和度が低い。特に、Sound 3 との組み合わせ条件は最も調和度が低い ($M = 0.5, SD = 0.9$)。それは、各文字の提示タイミングと効果音の呈示タイミングが時間的に同期していないためである。Sound 3 は、音刺激の中で最も短い持続時間 (266 ms) を持っており、映像の出現より 500 ms 遅れて鳴りはじまる。言い換えると、このことは音と映像の主観的調和の要因として時間構造の類似による構

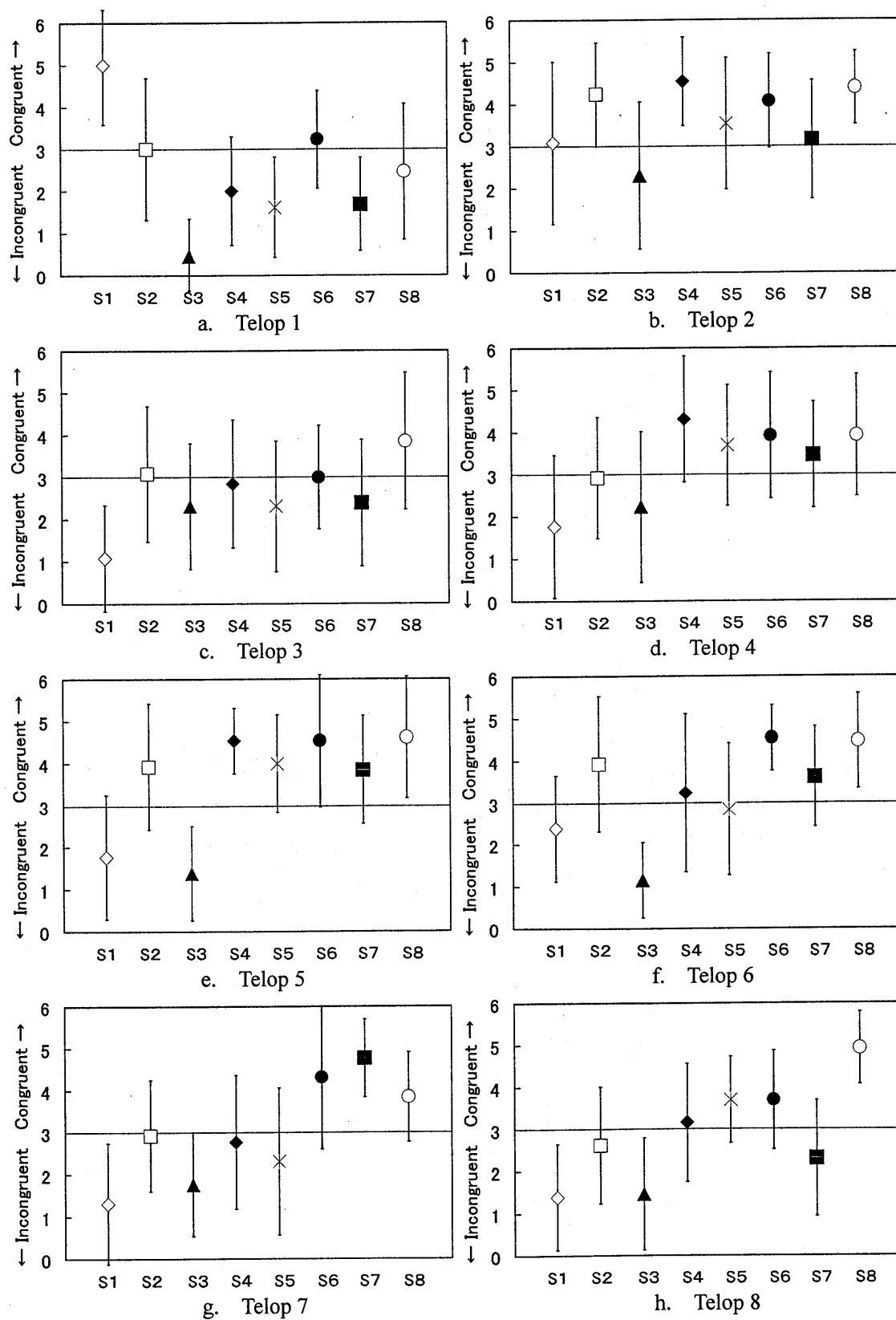


図2.5 実験1の映像刺激ごとに各効果音との組み合わせの平均調和評定値. Sound(S).
(エラーバーは標準偏差を表す)

造的調和の影響が大きいことを示唆するものである。

Telop 8 と Sound 8 のオリジナル条件は、全視聴覚刺激の中で2番目に調和度の高い組み合わせである。Telop 8 に対しては、すべての組み合わせ条件の調和度が低い。内観報告によると、「中洲の女」というテロップのことばや筆字からのイメージ（印象）と和太鼓の効果音がともに「和風」で調和を感じたという。テロップの言語情報の印象や書体の印象と効果音の印象の類似による意味的調和からも、高い主観的調和が得られることを示唆する。なお、この組み合わせではテロップと効果音の呈示のタイミングも同期していた。

全視聴覚刺激の中で3番目に調和度の高かった組み合わせは、Telop 7 と Sound 7 のオリジナル条件である。連続的に明るさが増加するテロップに対し、図 2.6 の楽譜に示すように Sound 7 はゆっくりと音高が上昇しており、音と映像の変化パターンの類似による構造的調和の影響が大きいと考えられる。一方、Telop 7 に対する組み合わせ条件ではどの刺激も調和度が低いが、Telop 7 と Sound 6 の組み合わせ条件はやや調和度が高い。Sound 6 は、高い周波数成分を持つ短音が不規則的な立ち上がりをみせながら複雑に変化するパターンである。テロップと効果音の変化パターンの類似による構造的調和においては、一つの映像に複数の音が合うかもしれない。



図 2.6 実験 1 の Sound 6 の楽譜

Telop 6 と Sound 6 のオリジナル条件は、Telop 6 に対する組み合わせ条件と比較して、相対的に調和度が高いが、Telop 6 と Sound 8 の組み合わせ条件の場合は同程度に調和度が高い。Telop 6 は、拡大しながら画面の前面に押し出してくる奥行き感のある映像である。Sound 6 は、短音が不規則的な立ち上がりをみせながら複雑に変化するパターンである。Sound 8 は、2連打の和太鼓の音である。

Telop 4 と Sound 4 のオリジナル条件も、Telop 4 に対する組み合わせ条件と比較して、相対的に調和度が高い。Telop 4 は、三つの変化パターンが組み合わせられたもので

ある。まず、奥行き方向の軸に沿って画面外から画面の中央に飛んできて停止する。そして、テロップの色が緑から黄色に変化しながら光った後、再び緑色に変わってゆくと同時にゆっくりと前の方へ拡大してくる。Sound 4 は、音が出はじめてから 210 ms を境に周波数成分が高域と低域に分けられている。映像が画面外から飛んできて停止するときに合わせて強く鳴り響く音が、音と映像の主観的調和をもたらしたと考えられる。

Telop 2 は、画面の中央でテロップが 1 回ゆっくりと光ってから消えるパターンである。Telop 2 の場合は Sound 2, Sound 4, Sound 6, Sound 8 との組み合わせで相対的に調和度が高く、これら以外の音刺激との組み合わせでは調和度が低い。調和する効果音においては、映像が明るく光りはじめる時点と対応する時点に音のエネルギーが集中しており、映像の明るさが下降するにつれて音の振幅エンベロープ（音量）も減衰している。このことから、音と映像の強さの変化パターンの類似による構造的調和の効果が大きいと考えられる。また、Telop 2 は、Telop 1 と言語情報、書体、文字の色は同じであり、持続時間は僅かな差（200 ms）を持つが、変化パターンはまったく異なる。よって、テロップの提示パターンが異なると調和する音も違うことが分かる。その上、Telop 1 のように時間構造の類似による構造的調和の場合は調和する映像と音の対応が一对一であるのに対し、Telop 2 のように変化パターンの類似による構造的調和の場合は映像と音の対応が一对多である可能性がみられた。

Telop 5 の場合も、Sound 4, Sound 5, Sound 6, Sound 8 のように複数の音刺激との組み合わせで相対的に調和度が高く、他の音刺激との組み合わせでは調和度が低い。Telop 5 のように、出現して消える単調なパターンは、Sound 4, Sound 5, Sound 8 のようにアクセント感のある音との組み合わせが調和する傾向があった。一方、Sound 1, Sound 3 との組み合わせ条件のように音と映像の時間構造が同期しない場合は調和度が低下する。

例外的に、Telop 3 はすべての音刺激との組み合わせで調和しないと評定された。Telop 3 と Sound 3 のオリジナル条件でも調和度が低い。Sound 3 は、音の開始部と終止部に音声が入っていたため、編集で取り除いた。それが原因で、Sound 3 とのすべての組み合わせは、映像の出現部と音の開始部の間（SOA）に 500 ms の時間ずれが生じ、映像の消滅部と音の終止部も同期していない。その結果、制作者が調和するように組み合わせたオリジナル条件も調和しなかった。この様に、視聴覚刺激の時間構造のずれは、

音と映像の主観的調和への影響が大きい。それは、音と映像が時間とともに変化するという特徴を持つからであると考えられる。

この様に、組み合わせ条件と比較して制作者が意図したオリジナル条件の方が、一般的に調和していると評定された。この傾向は、映像作品を用いた Iwamiya (1994), Lipscomb and Kendall (1994) の先行研究の結果と同様であり、制作者は視聴者の反応を想像しながらテロップと効果音を適切に組み合わせていることがうかがえた。

事後比較として Tukey (HSD) 法による多重比較を行い、映像刺激ごとにオリジナル条件と組み合わせ条件の水準間の調和度の相違を検定した。Telop 1 の場合、オリジナル条件と組み合わせ条件の間の調和感に、すべて統計的に有意な差が認められた ($p < .01$)。Telop 2 の場合、オリジナル条件と Sound 3 との組み合わせ条件のみの調和感に、統計的に有意な差が認められた ($p < .05$)。Telop 3 の場合、オリジナル条件と組み合わせ条件との調和感に、すべて統計的に有意な差はなかった ($p > .1$)。Telop 4 の場合、オリジナル条件と Sound 1, Sound 3 との組み合わせ条件のみの調和感に、統計的に有意な差が認められた ($p < .01$)。Telop 5 の場合も、オリジナル条件と Sound 1, Sound 3 との組み合わせ条件のみの調和感に、統計的に有意な差が認められた ($p < .01$)。Telop 6 の場合、オリジナル条件と Sound 1, Sound 3, Sound 5 との組み合わせ条件のみの調和感に、統計的に有意な差が認められた ($p < .01, p < .05$)。Telop 7 の場合、オリジナル条件と Sound 6, Sound 8 との組み合わせ条件の調和感には統計的に有意な差はないが ($p > .1$)、他の組み合わせ条件との調和感には統計的に有意な差が認められた ($p < .01$)。Telop 8 の場合も、オリジナル条件と Sound 5, Sound 6 との組み合わせ条件の調和感には統計的に有意な差はないが ($p > .1$)、他の組み合わせ条件との調和感には統計的に有意な差が認められた ($p < .01$)。

さらに、テロップと効果音の主観的調和と視聴覚情報の印象との関係を探るために、図 2.3 と図 2.4 の各視聴覚刺激の「評価/力動」主成分得点、「活動」主成分得点と図 2.5 の平均調和評定値との相関係数を求めた。その結果を示した表 2.4 より、統計的に有意であった値を見ると、平均調和評定値と活動主成分得点の間には弱い共変関係があるが、平均調和評定値と評価/力動主成分得点の間には強い共変関係が認められた。このことは、テロップと効果音の主観的調和度が高いと、視聴覚情報の総合的評価も高まることを意味する。岩宮 (2000) の研究では、音と映像が調和していない場合より調和すると

表 2.4 実験 1 の視聴覚刺激の印象と調和の相関係数

視聴覚刺激の調和	視聴覚刺激の印象	
	評価/力動主成分得点	活動主成分得点
平均調和評定値	.74**	-.28*

* 5% ** 1% 水準で有意

きに互いの効果を高め合う「協合現象」が生じやすいという結果が得られており、音と映像が調和している場合に映像作品の全体の評価が高まる効果は、協合現象と関連があると考えられる。

本実験では、通常のテレビ番組で多用されているテロップと効果音を実験刺激として用い、任意にテロップと効果音を組み合わせることによって、どの様に調和感が変化し、それがテロップの総合的評価にどのような影響を及ぼすのかを SD 法による印象評定実験を通して検討した。実験の結果、やはり制作者が意図したテロップと効果音の組み合わせ(オリジナル条件)の方が、制作者の意図に基づかない組み合わせ条件より調和した。こうしたテロップと効果音の主観的調和の要因は、構造的側面として両者の間の「時間構造」と「変化パターン」の類似が考えられた。また、テロップと効果音の構造的調和が音と映像の主観的調和を上昇させ、視聴覚情報の評価を高めるとの傾向が得られた。内観報告にも、テロップと効果音が「調和(同期)」していないと映像作品として「良い」とは思えなかったという意見が多数あり、被験者が音と映像が組み合わせられる際の評価の判断基準として「調和」を常に意識したことが分かる。さらに、内観報告からはテロップの書体の印象と効果音の印象、テロップのことばの印象と効果音の印象の類似のような意味的側面も、音と映像の主観的調和を高める要因として挙げられた。

ただし、本実験では、実際のテレビ番組のテロップと効果音を実験刺激として用いていたため、各種の要因が交錯しており、音と映像の構造的調和の影響を独立した形で確認することはできなかった。次節の実験 2 では、テロップと効果音の構造的調和の影響を独立して検証できるように、視聴覚刺激を系統的に作成して印象評定実験を行った。

2.3 実験2: 合成したテロップと効果音を用いた実験

2.3.1 はじめに

2.2節の実験1によって、テロップと効果音の構造的調和が音と映像の主観的調和を上昇させ、視聴覚情報の評価を高めるとの傾向が得られた。特に、テロップと効果音の時間構造の同期といった構造的調和の効果が多くみられた。なお、構造的調和には音と映像の時間構造の類似によるものと変化パターンの類似によるものがあることが示唆された。しかし、実際に放映されているテロップを用いたため、刺激の中には意味的調和の効果を避けられなかったものもあった。そこで、本節では実験条件を次の2.3.2項のように統制することによって、実験1で音と映像の構造的調和の要因として取り上げられた時間構造の同期と変化パターンの一致の効果を明らかにすることを目的とする。

また、視聴覚刺激の主観的調和と印象への音刺激と映像刺激の印象の影響も同時に検討する。

2.3.2 実験方法

まず、テロップと効果音の構造的調和の効果を検討するためには、両者の意味的側面の要因を排除した実験刺激が必要となる。そこで、テロップと効果音を合成し、できる限り構造的調和との対応関係のみが検討できるように実験条件を統制した。

テロップの変化パターンとしては、実験1に用いた刺激を参考にした7種類と新たな3種類を加え(Telop 2, Telop 7, Telop 10), 10種類を用いた。なお, Telop 1, Telop 2, Telop 4, Telop 5, Telop 6, Telop 7, Telop 10の表現技法は、文字の動的表示について検討した森本, 鳥丸, 西村(1992)も一般的に利用頻度の高いものと考えており、テロップ研究の刺激として妥当であると考えられる。

映像刺激は、文字が意味を持つとそれ自体が映像の印象に何らかの影響を及ぼす可能性があるため、7個の子音字のみを組み合わせた無意味な綴りを用いた。表記形態についても、72ptのアルファベットの小文字(gksmz dq)のみとし、文字の種類に統一感を与えた。こうした文字の種類の違いには、認知を助ける効果もある(Coltheart & Freeman, 1974)。また、文字の書体に強い特徴があるとそれも何らかの印象を生じさせる可能性があるため、一般性のあるTimes New Romanを選んだ。画面上に呈示した際の実寸は、

縦 (ascender line から descender line まで) 22×横 76 (mm) であった。配色は、被験者が文字の色と背景から何らかの印象を受けられると実験の結果に影響を及ぼすためにモノトーンとし、文字の色を白に、背景色を黒とした。こうした明度対比は、刺激を強めて対象の認知を助ける効果がある (長田, 長谷川, 福田, 安田, 吉田, 1987)。この様に、映像刺激はテロップの意味的側面の要因をできる限り排除させた。そして、テロップの呈示時間をすべて 2000 ms に設定し、時間構造と変化パターンを操作した。

映像刺激は Adobe 社の Photoshop 6.0j と After Effects 5.0 を用い、frame size を 640×480 (pixel), frame rate を 30 fps (frame per second) にし、key frame 法により動画表現 (animation) した。frame size は、NTSC カラー方式テレビのアスペクト比 (4:3) と同様である。図 2.7 に、実験 2 の各映像刺激を示す。

音刺激は、図 2.7 の映像刺激の時間構造と変化パターンを参考にし、純音を用いて図 2.8 に示すように 10 種類を合成した。音刺激の作成に純音を用いた理由は、現実音や楽器音などを使うと音の持つ意味が実験結果に何らかの影響を及ぼす可能性があり、音刺激の意味的側面の要因が排除できないからである。また、ノイズの場合は周波数との対応関係が分かり難いし、複合音の場合は倍音構造による音色の影響が考えられる。

すべての音刺激は、Syntrillium 社の Cool Edit Pro を用いて作成し、サンプリング周波数 44.1 kHz, 量子化ビット数 16 bit の WAVE ファイルに書き出した。刺激作成に用いた純音は、Sound 1 のように基本的には 1000 Hz, 2000 ms であり、すべての音刺激の持続時間には 10 ms の立ち上がりと立下りが含まれている。一方、音の時間構造を操作した Sound 2, Sound 3, Sound 4 は、それぞれに 1800 ms, 1120 ms, 200 ms の持続時間を持つ。また、周波数を操作した Sound 6 は、音高が 987.4 Hz から 1409 Hz まで連続的に上昇するグライド音 (上昇音高) である。Sound 7 は、基本周波数 1000 Hz, 変調周波数 1 Hz, 変調指数 100 の周波数変調音 (FM 音) である。Sound 10 は、B4 (987.4 Hz) から F#6 (2960 Hz) の間の 20 音を、まとまり感のないように音高をランダムに配置したものである。各音刺激の呈示音圧レベル (等価騒音レベル) は、実験 1 と同様の方法で設定した。その結果、Sound 1, 84 dB; Sound 2, 82 dB; Sound 3, 82 dB; Sound 4, 74 dB; Sound 5, 82 dB; Sound 6, 83 dB; Sound 7, 84 dB; Sound 8, 79 dB; Sound 9, 82 dB; Sound 10, 85 dB であった。

gksmzdq

(0 ms) ... (2000 ms)

a. Telop 1: Still (Steady state pattern)

gksmzdq

gksmzdq

gksmzdq

(0 ms) (360 ms) (720 ms) (1080 ms) (1440 ms) (2000 ms)

b. Telop 2: Twinkling

g

gk

gks

gksm

gksmz

(0 ms) (160 ms) (320 ms) (480 ms) (640 ms) (800 ms) ...

gksmzd

gksmzdq

..... (960 ms) ... (2000 ms)

c. Telop 3: Consecutively appearing pattern such as a "typewriter"

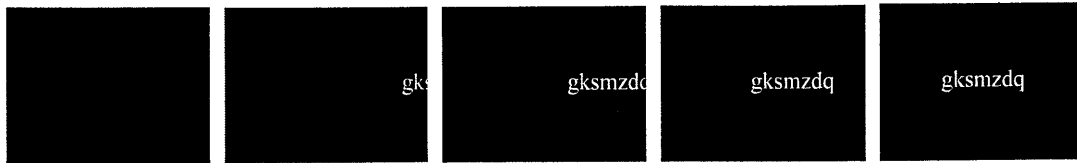
gk

gksmzdq

(0 ms) (33 ms) (200 ms) ... (2000 ms)

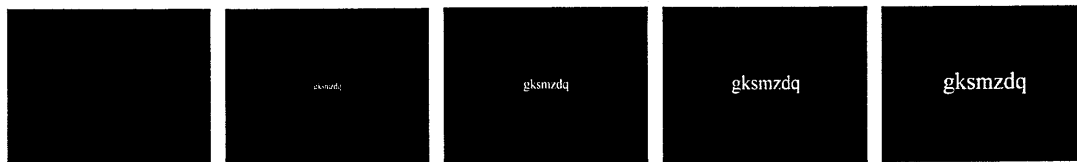
d. Telop 4: Fly in

CONTINUED ON NEXT PAGE



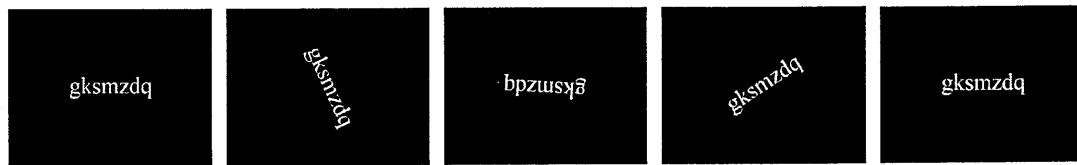
(0 ms).....(330 ms).....(1000 ms).....(1500 ms).....(2000 ms)

e. Telop 5: Slide in



(0 ms).....(500 ms).....(1000 ms).....(1500 ms).....(2000 ms)

f. Telop 6: Zoom in



(0 ms).....(330 ms).....(1000 ms).....(1660 ms).....(2000 ms)

g. Telop 7: Rotation



(0 ms).....(330 ms).....(660 ms).....(1000 ms).....(2000 ms)

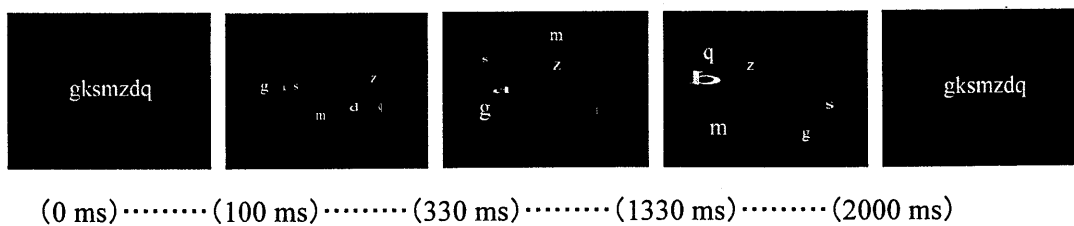
h. Telop 8: Fade in



(0 ms).....(330 ms).....(1000 ms).....(1660 ms).....(2000 ms)

i. Telop 9: Fade in & out

CONTINUED ON NEXT PAGE



j. Telop 10: Scrambling

図 2.7 実験 2 の各映像刺激の変化パターンとその特徴

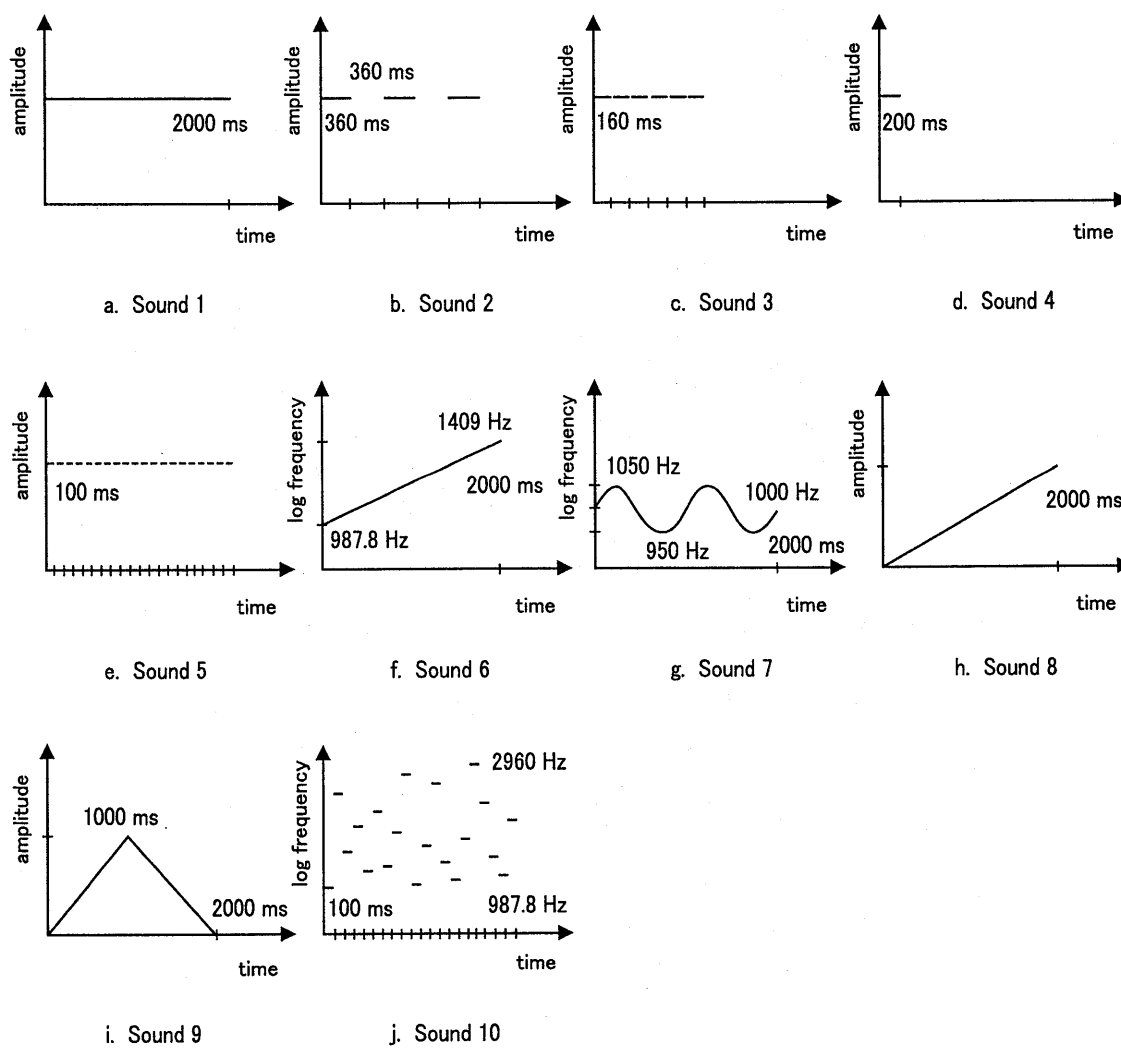


図 2.8 実験 2 の各音刺激の変化パターンとその特徴

印象評定尺度は、実験 1 に用いた尺度の中で 2 主成分に負荷の高い 1 対を除いた 9 対に、本実験に用いる刺激の特徴を考慮して 2 対を加え、11 対を用いた (表 2.5)。

被験者は、九州芸術工科大学の音響設計学科に在学する視力 (矯正視力を含む)、聴力ともに健常な日本人大学生および大学院生 20 名 (男性 17 名、女性 3 名) である。年齢は、22 歳から 29 歳 (平均年齢 23.6 歳) である。20 名の被験者のうち、12 名は実験 1 にも参加している。

実験においては、図 2.9 に示すように 最初の 0 ms から 1000 ms までは黒い画面の中央に直径 7 mm の白い小さい円を呈示し、被験者の注意を画面に向けさせる。画面と円の輝度は、輝度計 (MINOLTA LS-110) で測定した結果、それぞれ 0.11 cd/m^2 と 23.19 cd/m^2 であった。1000 ms から 2000 ms までは黒い背景のみの状態とする。その後、2000 ms から 4000 ms まで映像刺激を呈示する。音刺激単独の印象評定実験においては、音刺激は黒い画面のみの状態で被験者の両耳に呈示した。各刺激と 7 段階の両極の印象評定尺度は被験者ごとにランダムな順で呈示した。すべての評定は、実験 1 と同様にコンピュータの画面上で行った。被験者にはすべての刺激に対し、1 回呈示が終わってから評定するように教示した。また、評定が難しかった場合は何度も繰り返して視聴できるようにした。

実験としては、10 種類の音刺激単独の印象、10 種類の映像刺激単独の印象、そしてこれらの音刺激と映像刺激を組み合わせた 100 種類の視聴覚刺激の印象について SD 法による印象評定実験を行った。併せて、視聴覚刺激の印象評定実験では、7 段階の「調和していない—調和している」の尺度を用い、音と映像の主観的調和も評定させた。す

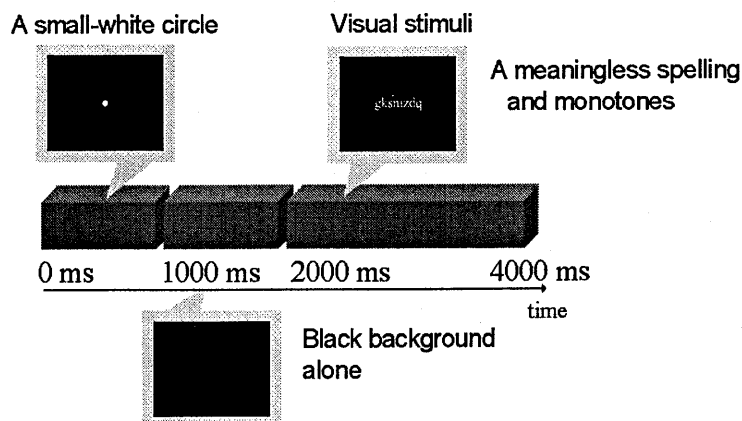


図 2.9 実験 2 の刺激呈示の概略

すべての実験において本試行に入る前に、被験者に実験環境と課題に慣れさせるため、1種類の刺激を用いて練習試行を行わせた。練習試行を含めて実験の補足説明に掛かった時間は約10分であり、被験者に暗所での刺激の見え方に支障がないことを確認した。視聴覚刺激の印象評定実験では、途中で3回の休憩を挟み、被験者の負担を減らした。実験環境と装置は、実験1と同様である。

2.3.3 結果と考察

視聴覚刺激の印象に関する評定尺度ごとに被験者間の平均評定値を求め、評定尺度を変量とした主成分分析を行った。表2.5に、バリマックス回転後の主成分係数行列を示す。

表 2.5 実験2の視聴覚刺激の印象評定実験より得られた主成分係数行列

印象評定尺度	主成分		
	力動	活動	評価
鈍いー鋭い	.958	.057	-.064
弱々しいー力強い	.897	.073	.304
ゆったりとしたースピード感のある	.857	.164	-.237
アクセント感のないーアクセント感のある	.815	.324	.284
暗いー明るい	.702	.523	.088
広がり感のないー広がり感のある	-.068	.910	.073
つまらないー面白い	.395	.861	.021
単調なー変化や変動のある	.370	.805	-.340
ありふれたーユニークな	.181	.655	-.651
分かり難いー分かり易い	.074	-.132	.942
悪いー良い	.305	.561	.762
寄与率 (%)	36.8	30.8	20.0

第1主成分は「鈍いー鋭い」「弱々しいー力強い」「ゆったりとしたースピード感のある」「アクセント感のないーアクセント感がある」などの尺度に負荷が高いことから「力動主成分」と解釈した。第2主成分は「広がり感のないー広がり感がある」「つまらないー面白い」「単調なー変動や変化のある」「ありふれたーユニークな」などの尺度に負荷が高いことから「活動主成分」と解釈した。本実験では、面白くてユニークな印象の刺激が動的であり、面白さとユニークさが力動感と結び付いたと考えられる。第3主成分は「分かり難いー分かり易い」「悪いー良い」などの尺度に負荷が高いことから「評価主成分」と解釈した。Osgood et al. (1957)によると、これらの3主成分は幅広いカテゴリーの概念 (concepts) から得られるものという。音と映像の相互作用に関する先行研究 (Lipscomb & Kendall, 1994; Marshall & Cohen, 1988; Sirius & Clarke, 1994; Tannenbaum, 1956) でも、この普遍的な3主成分が使われた。また、これらの3主成分に対する各視聴覚刺激の主成分得点を求め、それぞれを図2.10、図2.11、図2.12に示す。

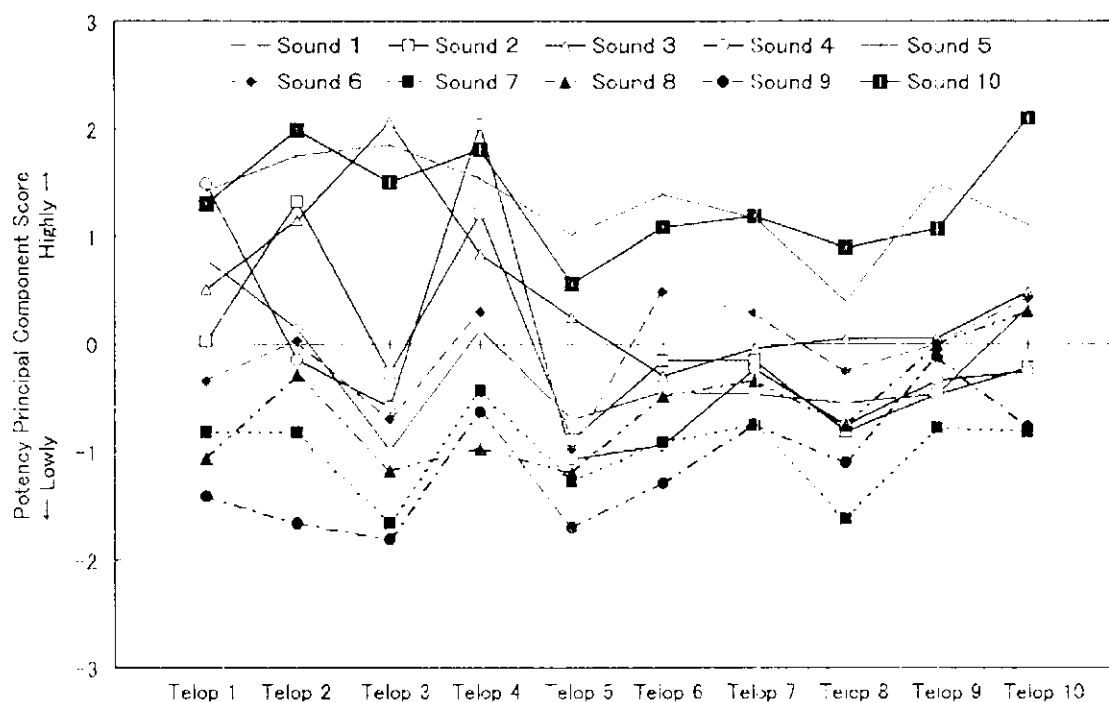


図2.10 実験2の各視聴覚刺激の力動主成分得点

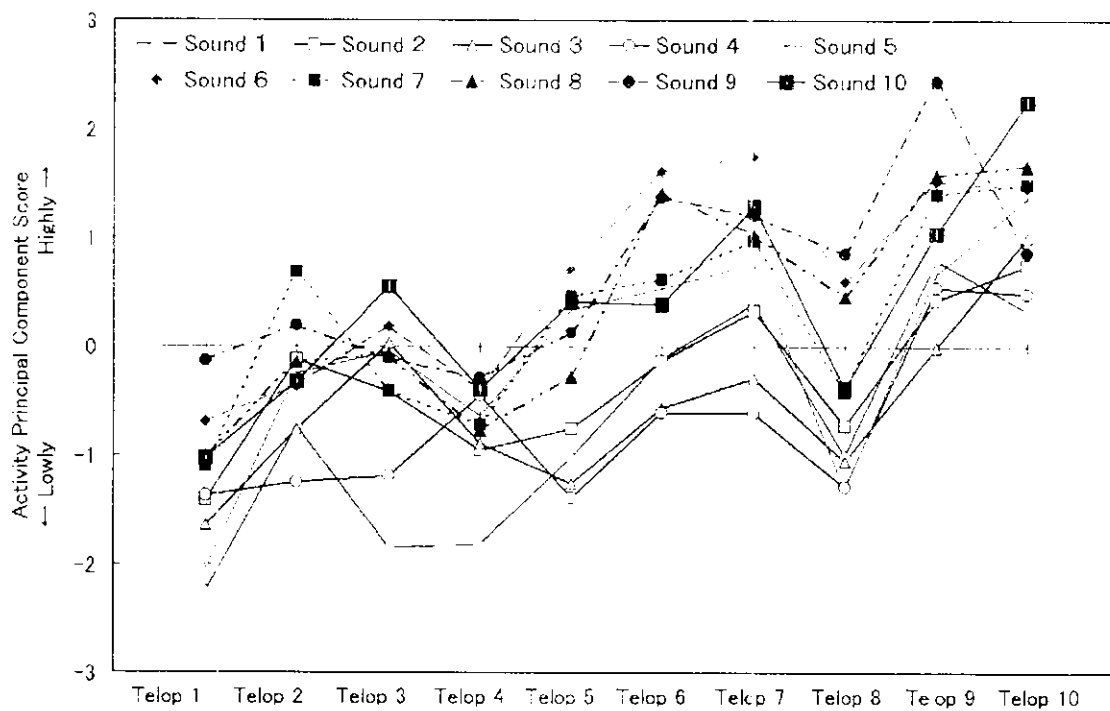


図 2.11 実験 2 の各視聴覚刺激の活動主成分得点

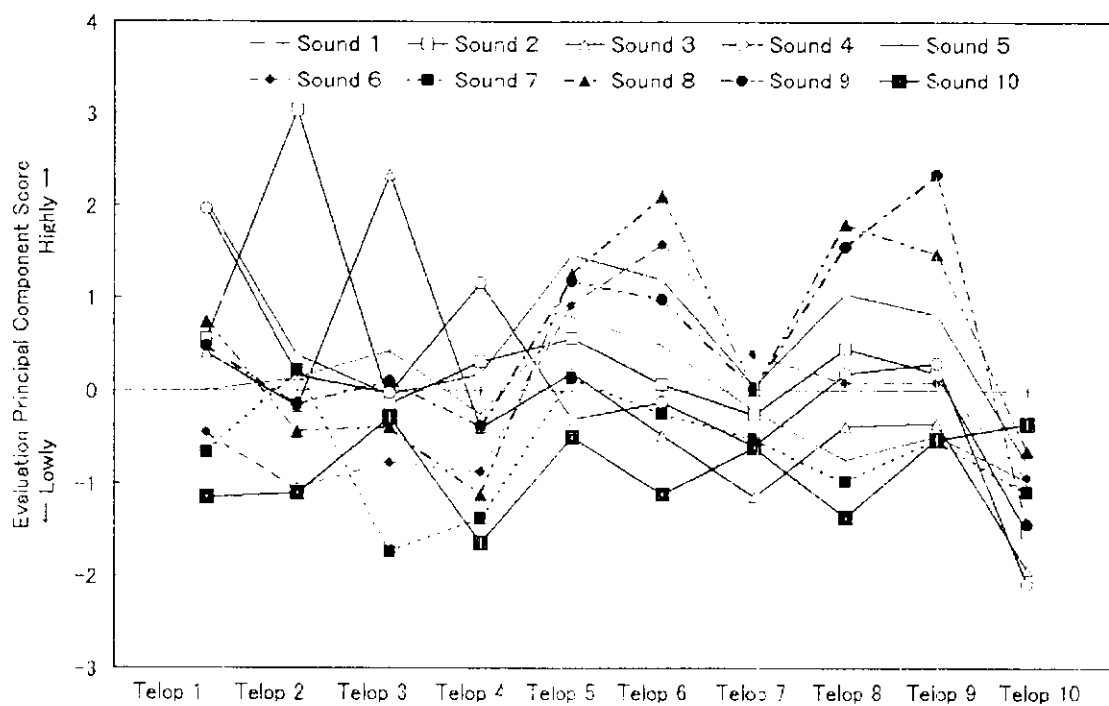


図 2.12 実験 2 の各視聴覚刺激の評価主成分得点

図 2.10 の視聴覚刺激の力動主成分得点では、Sound 5, Sound 7, Sound 8, Sound 9, Sound 10 のように各映像刺激との組み合わせで、極端な力動感の変化は見られず音刺激による影響が大きい傾向がある。一方、図 2.11 の視聴覚刺激の活動主成分得点を見ると、Telop 1, Telop 4, Telop 9, Telop 10 のように全般的にテロップに各効果音を組み合わせても活動感の変化が極端に変わることはなく、音刺激より映像刺激によって印象付けられる傾向がある。しかし、図 2.12 の視聴覚刺激の評価主成分得点は、Telop 2 と Sound 2, Telop 3 と Sound 3 の組み合わせのように、特定の映像刺激と音刺激が組み合わせられることによって高まる傾向がある。

さらに、これらの 3 主成分の主成分得点は、2.3.3.2 項の視聴覚刺激の印象に関する検討に用いられた。

2.3.3.1 テロップと効果音の時間構造と変化パターンが印象に与える影響

2.3.2 項で操作したテロップと効果音の時間構造と変化パターンから受ける印象を探るため、映像刺激単独と音刺激単独の印象評定実験より得られたデータをもとにプロフィールを作成した。図 2.13 には映像刺激単独の印象評定実験の結果を、図 2.14 には音刺激単独の印象評定実験の結果を示す。

図 2.13 の映像刺激の場合、出現して消える Telop 1 は「悪いー良い」「分かり難いー分かり易い」といった評価性次元に対する評定値は高いが、他の力動性次元と活動性次元に関わる尺度ではすべて評定値が低い。逆に、複雑に変化する Telop 10 の場合、力動性次元と活動性次元に関わる尺度ではすべて評定値が高いが、評価性次元に関わる尺度の評定値は低い。また、一文字ずつ呈示される Telop 3 と徐々に明るくなる Telop 8 の場合は評価性次元に対する評定値が高いが、回転する Telop 7 の場合は低い。この様に、テロップの変化が複雑でない方が文字を読み取りやすく（分かり易く）、良い印象を受けやすい。力動性次元に関わる尺度においては、相対的に Telop 4, Telop 10 の評定値が高く、Telop 5, Telop 8 の評定値は低い。「slide in」「fade in」のようにテロップがゆっくりと連続的に変化（移動）する場合より、「fly in」「scrambling」のようにアクセント感があったり複雑に変化したりする方が力動感のある印象になりやすい。活動次元に関わる尺度においては、相対的に Telop 7, Telop 9, Telop 10 の評定値が高く、特に Telop 1 の評定値は低い。出現して消える単調なテロップより、「rotation」

「fade in & out」「scrambling」のように連続的に変化したり複雑に変化したりする映像の方が活動感のある印象になりやすい。

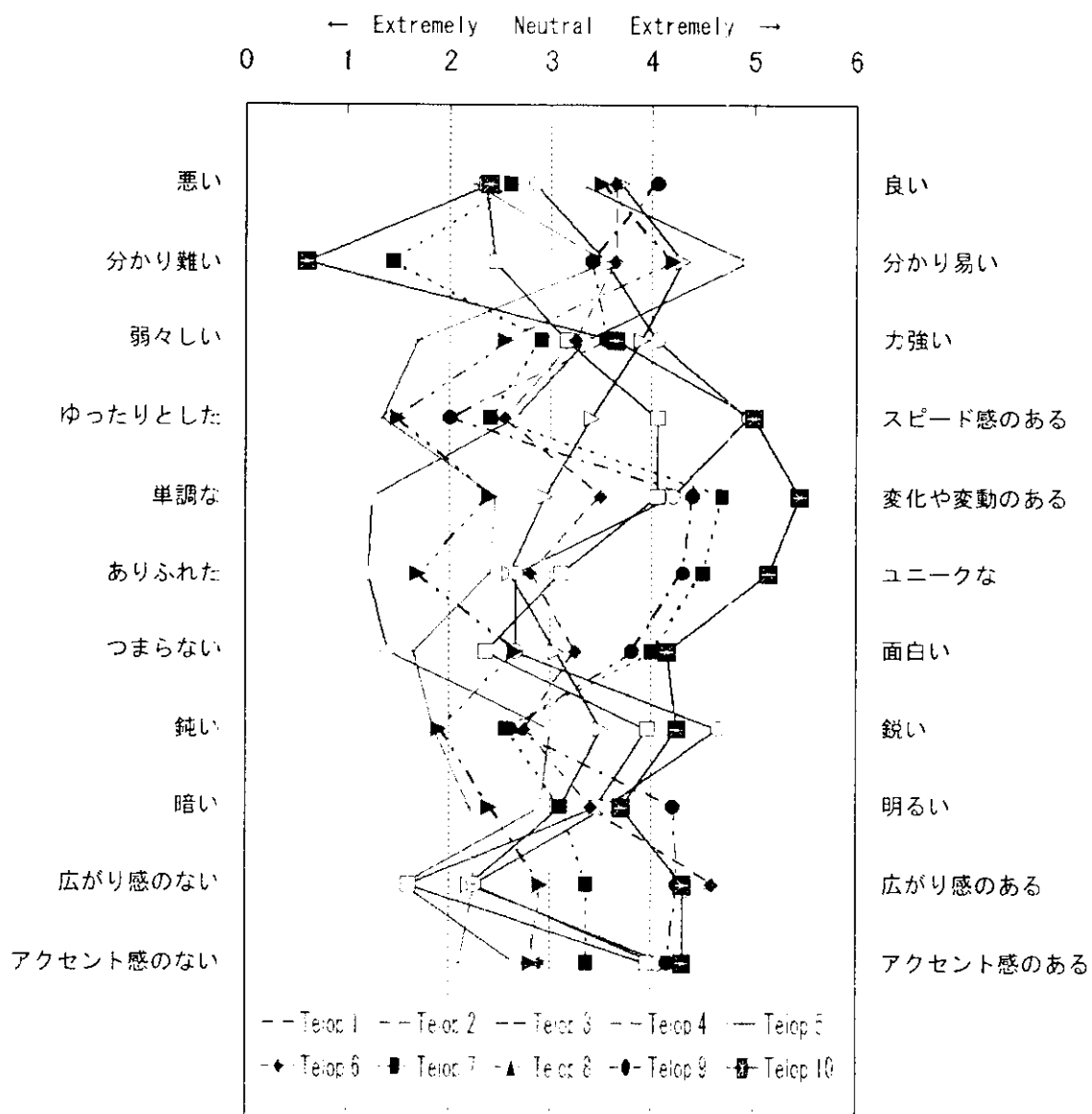


図 2.13 実験 2 の各映像刺激の質のグラフ

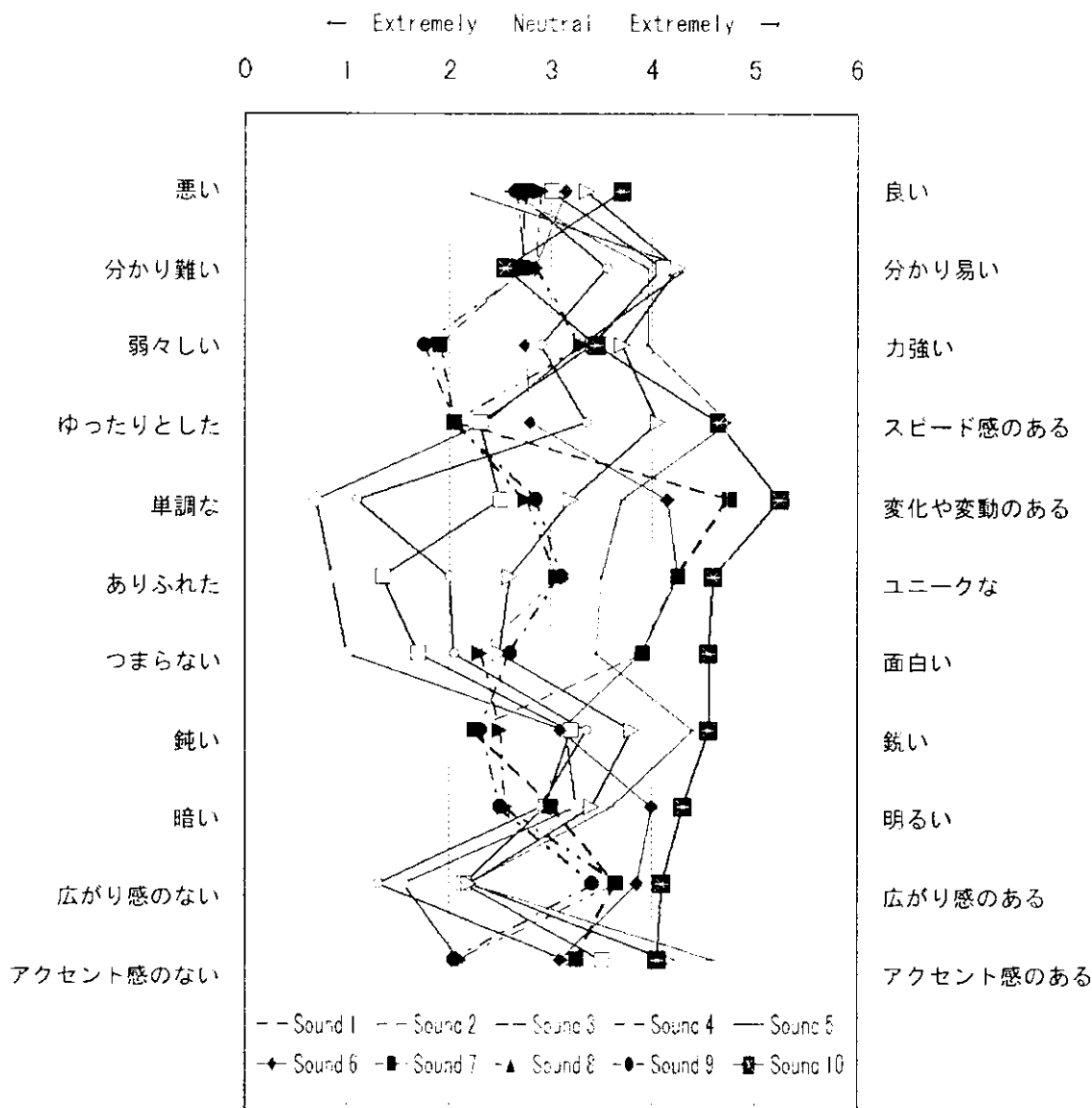


図 2.14 実験 2 の各音刺激のプロファイル

図 2.14 の音刺激の場合、評価性次元に関わる尺度においては、相対的に Sound 3 がやや評定値が高いが、その他の音ではほとんど差が見られない。それは、音刺激として純音を用いたためであると考えられる。活動性次元に関わる尺度においては、相対的に Sound 5, Sound 10 の評定値が高く、Sound 7, Sound 9 の評定値は低い。周波数変調音 (FM 音) や振幅エンベロープ (音量) が連続的に増大して減少する音のように連続的な変化より、断続的・複雑な変化を持つ音の方が活動感のある印象になりやすい。活動性次元

に関わる尺度においては、相対的に Sound 6, Sound 7, Sound 10 の評定値が高く、Sound 1, Sound 2, Sound 4 の評定値は低い。立ち上がって立ち下がる単調な音や刺激間時間間隔 (ISI: inter stimulus interval) の長い音より、連続的に音の高さや大きさが変化したり、複雑に音高が変化したりする音の方が活動感のある印象になりやすい。

2.3.3.2 視聴覚刺激の印象への音刺激の印象と映像刺激の印象の影響

まず、映像刺激と音刺激の印象評定尺度における平均評定値を用い、視聴覚刺激の主成分分析より得られた3主成分と対応するそれぞれの「合成得点」を算出するために主成分分析を行った。例えば、表 2.5 の視聴覚刺激の「力動」主成分に負荷の高かった「鈍いー鋭い」「弱々しいー力強い」「ゆったりとしたースピード感のある」「アクセント感のないーアクセント感のある」の尺度は、それぞれ映像刺激と音刺激の力動主成分に関わる合成得点を、「広がり感のないー広がり感のある」「つまらないー面白い」「単調なー変動や変化のある」「ありふれたーユニークな」の尺度は、それぞれ映像刺激と音刺激の活動主成分に関わる合成得点の算出に用いられた。同様に、視聴覚刺激の「評価」主成分に負荷の高かった尺度を用い、それぞれ映像刺激と音刺激の評価主成分に関わる合成得点を算出した。主成分分析の結果、映像刺激の印象に関する評定尺度の平均評定値は、それぞれ「力動 (84.8%)」「活動 (82.3%)」「評価 (80.0%)」の1主成分に集約された。なお、括弧内の値は寄与率を示す。また、音刺激の印象に関する評定尺度の平均評定値も、それぞれに「力動 (82.2%)」「活動 (88.6%)」「評価 (68.2%)」の1主成分に集約された。

視聴覚刺激の印象への映像刺激の印象と音刺激の印象の影響を調べるため、これらの映像刺激と音刺激の主成分得点と 2.3.3 項の視聴覚刺激の主成分得点を用いた。そして、視聴覚刺激の3主成分の主成分得点を従属変数、映像刺激と音刺激の3主成分の主成分得点を独立変数に重回帰分析を行った。図 2.15 の重回帰分析の結果を示したパス図には、映像刺激、音刺激、視聴覚刺激の「力動」「活動」「評価」主成分に対する重決定係数 (R^2) と標準偏回帰係数 (β) が示されている。

力動主成分と活動主成分の R^2 の値を見ると、それぞれ 1 に近くて独立変数全体が従属変数を予測/説明する程度が良い。また、 β の値を見ると、力動主成分の場合は映像

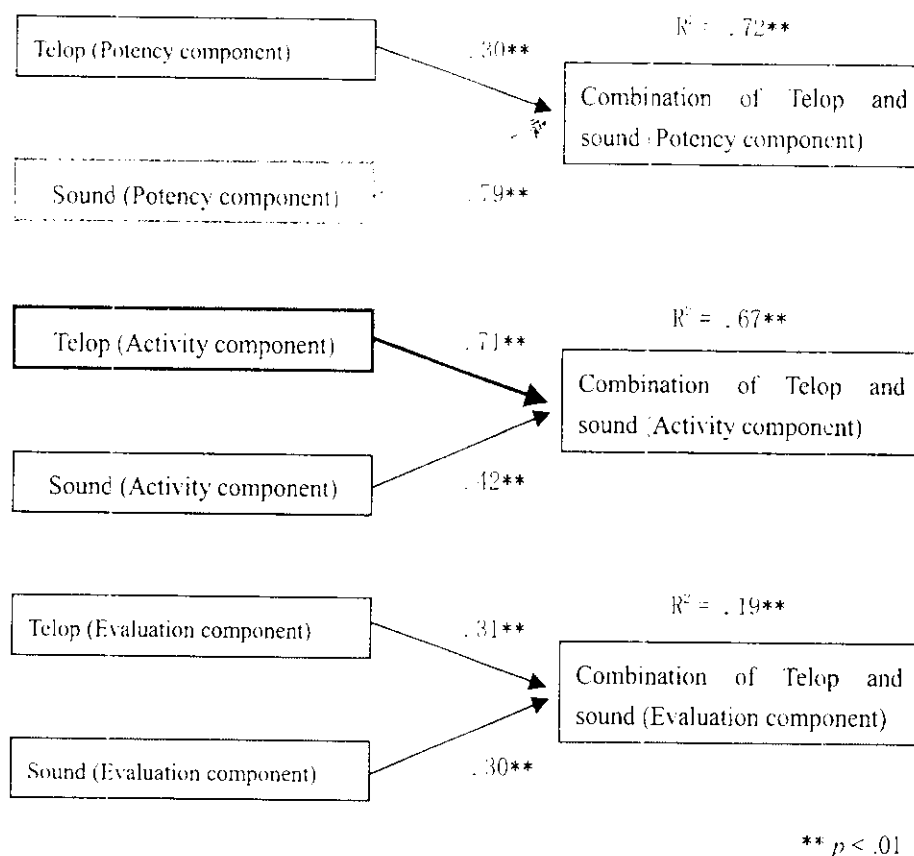


図 2.15 実験 2 の映像刺激、音刺激、視聴覚刺激の印象に関する重回帰分析の結果

刺激より音刺激の方が大きい。活動主成分の場合は音刺激より映像刺激の方が大きい。従って、テロップと効果音を組み合わせた際の力動感のレベルは、テロップの力動感よりも効果音の力動感からの影響が大きい。逆に、テロップと効果音を組み合わせた際の活動感のレベルは、効果音の活動感よりもテロップの活動感からの影響が大きい。Marshall and Cohen (1988) は、図 2.16 に示すように音楽の「力動性」と「活動性」次元が視聴覚情報の統合過程を経らずに映画の「力動性」と「活動性」次元を直ちに決定すると仮定し、これを「直接効果 (direct effect)」と呼んだ。Bolivar et al. (1994) は、映像メディアにおける背景音楽の直接効果を、音楽の意味が映像に付加されてその意味や解釈が付加された音楽によって系統的に変化することと考えた。本実験の結果、視聴覚刺激の「力動」の側面においては音の力動感による「直接効果」を実証した。しかし、視聴覚刺激の「活動」の側面においては Marshall and Cohen の仮定とは異なり、映像の活動感による「直接効果」が存在した。これは、他の感覚モダリティに比べて運

動情報の多くが視覚によって得られる（北島，1997）という運動知覚の特徴を考慮に入れば当然の結果である。

一方，図 2.15 の評価主成分の R^2 を見ると，統計的には有意であるが， R^2 の値が 0 に近くて独立変数全体が従属変数を予測/説明する程度が悪い。よって，テロップと効果音を組み合わせた際の評価のレベルは，テロップの評価からも効果音の評価からも影響は小さい。図 2.16 を見ると，Marshall and Cohen は映画の力動性次元と活動性次元においては音楽の意味が直接に影響するが，評価性次元の評価は音と映像の「比較照合器 (comparator)」によって決められる一種の「認知的調和 (cognitive congruency)」に依存すると仮定した。また，Lipscomb and Kendall (1994) も，映画の活動性/力動性は，映像に関係なく音楽の意味による影響が顕著であるが，評価に関しては Marshall

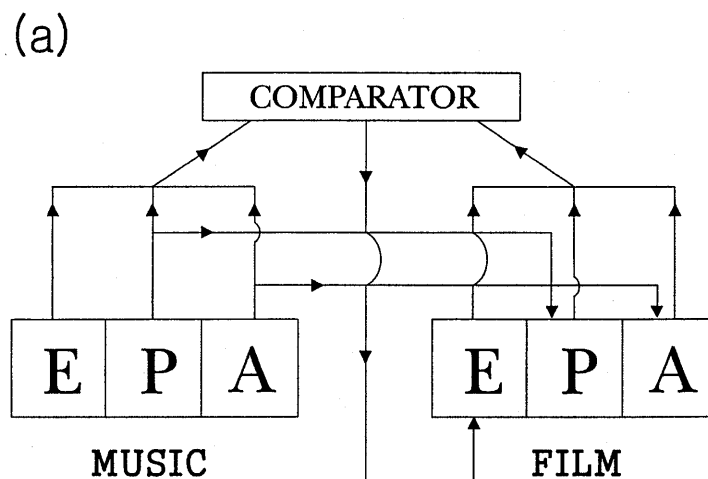


図 2.16 Marshall and Cohen (1988) の調和—連合モデル (a). Overall Evaluative (E), Potency (P), and Activity (A) meanings are shown for the music and film. Direct effects of the music are shown for the Potency and Activity levels. Evaluative levels depend upon the overall meaning from the three dimensions of both media. The degree of congruence found in the Comparator establishes the Evaluative meaning of the film and the music. Marshall and Cohen (1988) より転載.

表 2.6 実験 2 の視聴覚刺激の印象と調和の相関係数

視聴覚刺激の調和	視聴覚刺激の印象		
	力動主成分得点	活動主成分得点	評価主成分得点
平均調和評定値	.25*	.47**	.74**

* 5% ** 1% 水準で有意

and Cohen のように音と映像の組み合わせの適切さによって影響されると主張した。Iwamiya and Ozaki (2004), 岩宮ら (2005) は、視聴覚刺激の力動感と活動感は音によって強化されたが、評価においては音と映像の特定の変化パターンが組み合わせられるときに高まることを示した。この様に、視聴覚情報の評価の側面は、視覚情報と聴覚情報の間の調和の「認知的査定 (cognitive assessment)」に基づくものであると考えられる。

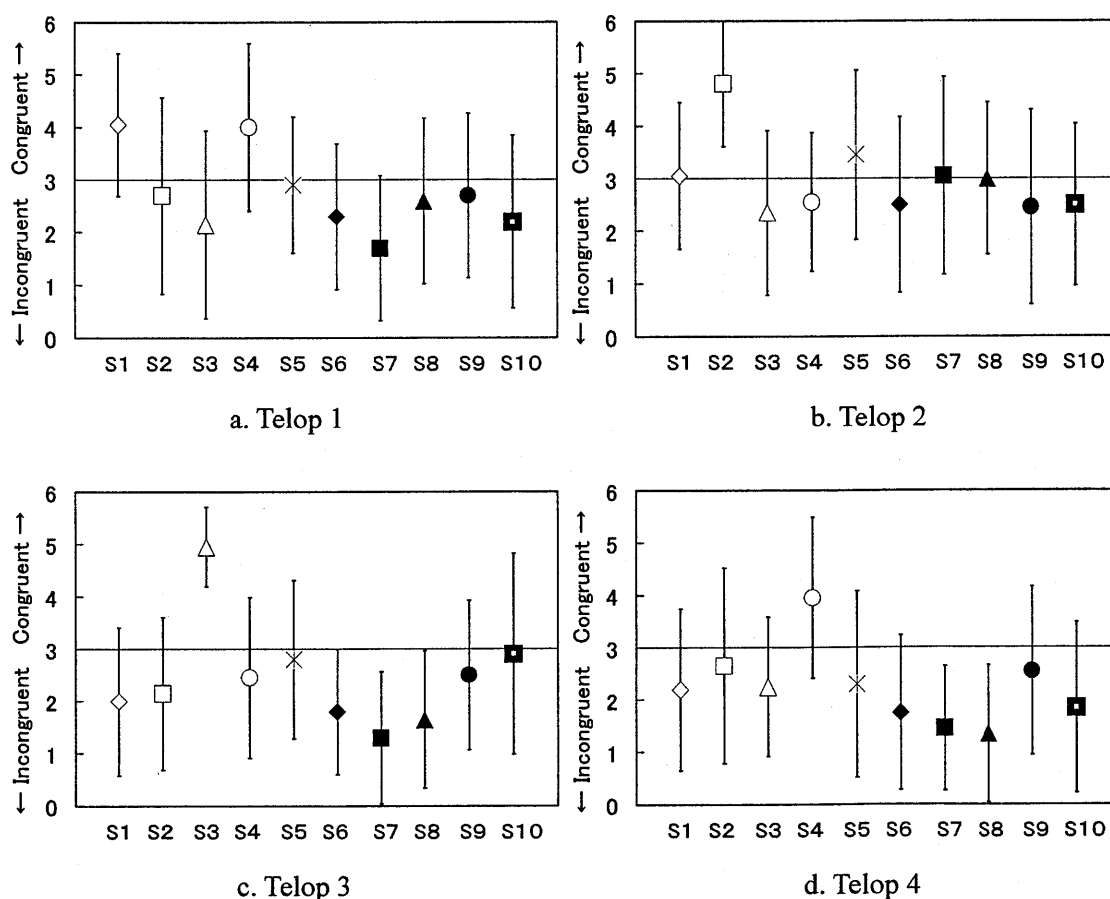
視聴覚刺激の主観的調和と印象との関連を調べるため、各視聴覚刺激の平均調和評定値と「力動」「活動」「評価」主成分得点との相関係数を求めた。その結果を示した表 2.6 より、平均調和評定値と「力動」「活動」主成分得点とは弱い共変関係があるが、平均調和評定値と「評価」主成分得点の間には強い共変関係があった。つまり、組み合わせた音と映像の調和度が高いと、視聴覚情報の評価も高まる。この様な傾向は、実験 1 でも得られている。

以上の結果をまとめると、視聴覚情報の力動感は聴覚情報より、活動感は視覚情報より直接に影響を受けるが、評価に関しては音と映像の組み合わせの主観的調和が重要である。なお、本研究によって視聴覚情報の評価、力動性、活動性の評定に関する Marshall and Cohen (1988) の仮説が実証された。

2.3.3.3 テロップと効果音の時間構造と変化パターンの構造的調和

本項では、視聴覚刺激の構造的調和の要因について検討する。視聴覚刺激の主観的調和に対する実験要因 (映像 10 水準×音 10 水準) の分析のために反復測定による二元配置の分散分析を行った。その結果、映像刺激の主効果 [$F_{(9,171)} = 6.36, p < .01$], 音刺激の主効果 [$F_{(9,171)} = 5.08, p < .01$], 映像刺激×音刺激の交互作用 [$F_{(81,1539)} = 7.45,$

$p < .01$] が、それぞれ統計的に有意であった。二つの要因間に交互作用が存在しており、各要因の水準間の差の検定はあまり意味を持たないため、水準ごとに一元配置の分散分析を行った (石村, 1992, 2002)。その結果, Telop 1 [$F_{(9, 171)} = 5.39, p < .01$], Telop 2 [$F_{(9, 171)} = 4.83, p < .01$], Telop 3 [$F_{(9, 171)} = 12.40, p < .01$], Telop 4 [$F_{(9, 171)} = 5.41, p < .01$], Telop 5 [$F_{(9, 171)} = 4.90, p < .01$], Telop 6 [$F_{(9, 171)} = 8.74, p < .01$], Telop 7 [$F_{(9, 171)} = 5.36, p < .01$], Telop 8 [$F_{(9, 171)} = 10.13, p < .01$], Telop 9 [$F_{(9, 171)} = 6.10, p < .01$], Telop 10 [$F_{(9, 171)} = 7.92, p < .01$] のように、各映像刺激における効果音の主効果がすべて統計的に有意であった。図 2.17 に、映像刺激ごとに各音刺激との組み合わせに対する平均調和評定値を示す。



CONTINUED ON NEXT PAGE

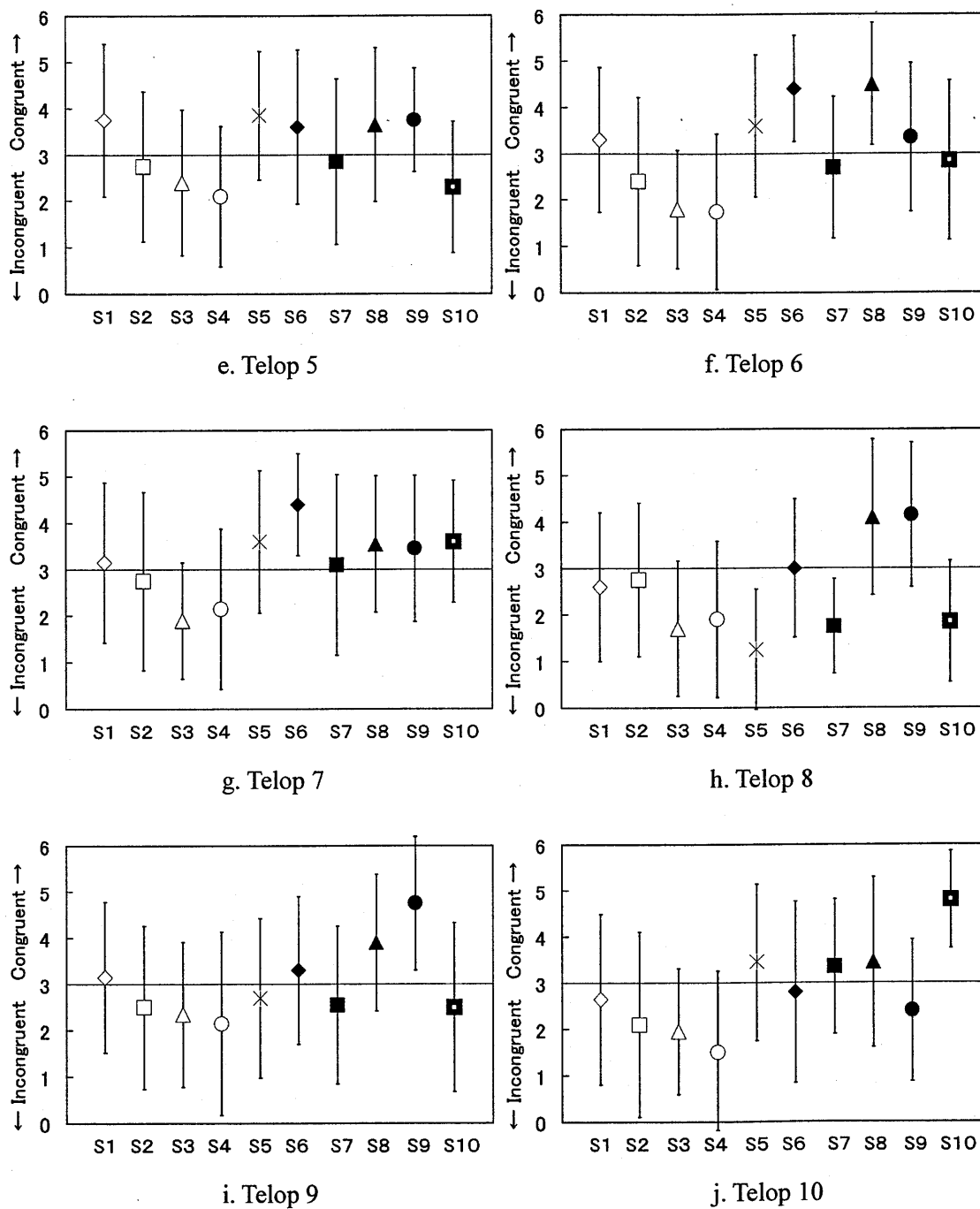


図 2.17 実験 2 の映像刺激ごとに各効果音との組み合わせの平均調和評定値. Sound(s).

(エラーバーは標準偏差を表す)

図 2.17 より、各映像刺激において Telop 1 は Sound 1, Telop 2 は Sound 2, Telop 3 は Sound 3, Telop 4 は Sound 4, Telop 5 は Sound 5, Telop 6 は Sound 8, Telop 7 は Sound 6, Telop 8 は Sound 9, Telop 9 は Sound 9, Telop 10 は Sound 10 との組み合わせで、他の音刺激との組み合わせより調和度が高い傾向がある。便宜的に、平均調和評定値が 4 以上の視聴覚刺激を調和度の高いものと定義すると、Telop 2, Telop 3, Telop 7, Telop 9, Telop 10 は、それぞれ Sound 2, Sound 3, Sound 6, Sound 9, Sound 10 との組み合わせで調和度が高い。これらの視聴覚刺激は、特定のテロップ・パターンとサウンド・パターンが組み合わせられており、高い調和感が得られた。一方、Telop 1 は Sound 1 と Sound 4, Telop 6 は Sound 6 と Sound 8, Telop 8 は Sound 8 と Sound 9 のように、特定のテロップ・パターンに複数のサウンド・パターンが調和した。以下は、音と映像の構造的調和を時間的側面における時間構造の類似と空間的/強さ（明るさ）の側面における変化パターンの類似から検討する。

Telop 1 は、実験 1 の Telop 5 を参考にして合成した。持続時間 2000 ms の出現して消える単調なテロップ・パターンに、持続時間 2000 ms の Sound 1（定常音）と 200 ms の Sound 4（単音）といった立ち上がって立ち下がる単調なサウンド・パターンが調和した。これらの視聴覚組み合わせは、映像の出現部と音の開始部が同期した。なお、Telop 1 と Sound 1 の場合は、映像の消滅部と音の終止部も同期している。この様に Telop 1 と Sound 1, Sound 4 の組み合わせから知覚された調和感は、音と映像の時間構造の類似による構造的調和に基づくものと考えられる。しかし、Telop 1 と Sound 4 の組み合わせは、映像の消滅部と音の終止部の間に 1800 ms のずれ（asynchrony）があったにもかかわらず調和していると判断された。これは、音と映像の時間構造の類似に基づく構造的調和は、両者の「onset（開始）」の同期が重要であることを示唆するものである。一方、Sound 7（FM 音）、Sound 6（上昇音高）のように Telop 1 と同じ時間構造を持っていても変化のあるサウンド・パターンとの組み合わせでは調和度が低い。よって、単調な変化の映像には単調な変化の音が調和感をもたらすと考えられる。また、Telop 1 は Sound 3（断続音）との組み合わせのように時間構造が合わない場合も調和度が低い。

Telop 2 は、360 ms 間隔で規則的に 3 回点滅する。Telop 2 と最も調和した Sound 2 は、360 ms の持続時間と刺激間時間間隔（ISI）を持つゆっくりとした「点滅音」である。この視聴覚組み合わせでは、映像の出現部と音の開始部、映像の消滅部と音の終止

部が同期した。よって、音と映像の時間構造の類似による構造的調和が要因で、調和度が高いと考えられる。一方、Telop 2 と時間構造が同期していない Sound 3 (断続音)、Sound 4 (単音) との組み合わせでは、調和度が低い。Sound 3 は、持続時間 160 ms の七つの単音で構成された「断続音」である。Sound 4 は、持続時間 200 ms の「単音」である。さらに、Telop 2 は、Sound 6 の音高が上昇するグライド音 (上昇音高)、Sound 9 の振幅エンベロープ (音量) が連続的に増大して減少する音 (増大減少音)、Sound 10 の不規則な音高の進行 (random 音高) のような変化パターンを操作した音とも調和度が低い。

Telop 3 は、実験 1 の Telop 1 を参考にして合成した。Telop 3 は、160 ms の持続時間を持つ 7 個の文字が「タイプライター」のように連続的に呈示される。Telop 3 と Sound 3 (断続音) は、一番調和度の高い組み合わせである。この視聴覚組み合わせは本実験の Telop 2 と Sound 2 の組み合わせと同様に、映像の出現部と音の開始部、映像の消滅部と音の終止部が同期した。よって、音と映像の時間構造の類似による構造的調和は視聴覚情報の主観的調和を高める効果がある。しかし、Telop 3 と時間構造が類似していない Sound 1 (定常音)、Sound 2 (点滅音)、Sound 4 (単音) との組み合わせは、調和度が低い。さらに、Telop 3 は、Sound 6 (上昇音高)、Sound 7 (FM 音)、音量が徐々に増大する Sound 8 (増大音) のような変化パターンを操作した音とも調和度が低い。

Telop 4 は、実験 1 の Telop 4 を参考にして合成した。Telop 4 は、0 ms から 200 ms にかけて 55 cm/s の速さで右側の画面外から画面の中央に素早く飛んできた後、2000 ms まで変化しないパターンである。Telop 4 は Sound 4 (単音) との組み合わせで調和度が高い。この視聴覚組み合わせは、映像の動きの開始部と音の開始部、映像の動きの終止部と音の終止部が同期している。しかし、本実験の Telop 1 と Sound 4 の組み合わせのように、映像の消滅部と音の終止部の間には 1800 ms の同期ずれがある。すなわち、本実験の Telop 1 と Sound 4 の組み合わせと同様に音と映像の時間構造の類似による構造的調和は、両者の「onset」の同期によって強く影響された。一方、Telop 4 と Sound 6 (上昇音高)、Sound 7 (FM 音)、Sound 8 (増大音)、Sound 10 (random 音高) のような変化パターンを持つ音との組み合わせは、調和度が低い。また、Telop 4 は Sound 1 (定常音)、Sound 3 (断続音) のような時間構造が同期していない音とも調和度が低い。

以上の Telop 1, Telop 2, Telop 3, Telop 4 と調和するサウンド・パターンは、い

ずれもその映像と類似した時間構造を持っていた。特に、音と映像の「onset」が同期する場合、その「offset」が同期しなくても高い調和度が得られた。一方、音と映像の時間構造が同期していないと調和度は低下した (e. g., Telop 2 と Sound 3, Telop 3 と Sound 2)。また、Sound 6 (上昇音高), Sound 7 (FM 音), Sound 8 (増大音) のように変化パターンを操作した音とも調和しなかった。これらの結果は、音と映像の時間構造の類似による構造的調和が視聴覚情報の主観的調和に貢献することを示すものである。本実験から得られた結果は、音と映像の時間的なアクセント構造の同期について検討した Lipscomb (1996, 1999) などの研究の結果とも一致しており、ミッキー・マウシングの効果を確認したものである。

Telop 5 は、実験 1 の Telop 3 を参考にして合成した。Telop 5 は、0 ms から 2000 ms にかけて 5.5 cm/s の速さで右側の画面外から画面の中央にゆっくりと滑り込んでくる。Telop 5 は Sound 1 (定常音), 持続時間 100 ms の単音 20 個が鳴る Sound 5 (単音列), Sound 6 (上昇音高), Sound 8 (増大音), Sound 9 (増大減少音) との組み合わせで相対的に調和度が高い。左方向へ水平に動くテロップ・パターンには、単調な変化を持つ多様なサウンド・パターンが調和した。また、Sound 1 (定常音) との組み合わせのように映像の動きの開始部と音の開始部、映像の動きの終止部と音の終止部の同期による影響もみられる。また、Sound 5 (単音列) との組み合わせの場合、物理的に連続的な時間構造でないにもかかわらず隣接する音同士の刺激間時間間隔 (ISI) が非常に短いため、調和していると評定されたと考えられる。一方、Telop 5 と Sound 3 (断続音), Sound 4 (単音), Sound 10 (random 音高) との組み合わせでは、調和度が低い。この様に、Telop 5 は時間構造の非同期や不規則な変化のない限り、各種の音の変化パターンと調和すると評定された。特に、水平面において左方向へ移動する映像と音の高さの上昇と大きさの増大が対応することは興味深い。

Telop 6 は、実験 1 の Telop 6 を参考にして合成しており、最大に変化した文字列の大きさは 16.72 cm² の内に収まる。Telop 6 は、Sound 6 (上昇音高), Sound 8 (増大音) との組み合わせで調和度が高い。これらの視聴覚組み合わせは、奥の方から前の方へ徐々に拡大して (近づいて) くるテロップ・パターンに、連続的に音高が上昇したり、徐々に音量が増大したりする変化を持つサウンド・パターンが調和した。有田, 岩宮 (2004), 岩宮ら (2003, 2005) の研究でも、拡大する映像と上昇形の音列が調和する

結果が得られた。さらに、岩宮ら (2005) の研究では、映像の拡大と音量の増大も相性が良かった。この様に、映像の変化パターンと音の変化パターンの構造的調和には「多義性」が存在する。本実験でも同様の結果が得られており、音と映像の空間的側面における変化パターンの類似による構造的調和が確認できた。その原因としては、ドップラー効果 (Doppler effect) の連想が考えられる。Neuhoff and McBeath (1996) は、音源が近づくとつれて音高が上昇するという一般的にいわれていることが実際に生じている (i. e., *Doppler illusion*) ことを確認した。また、音の強さの場合、自由音場で点音源 (or 線音源) のラウドネスは逆二乗則によって増加/減衰する。この様に映像の奥行きの変化は、音の高さと大きさの変化と対応すると考えられる。一方、Telop 6 と Sound 2 (点減音), Sound 3 (断続音), Sound 4 (単音) との組み合わせは、調和度が低い。徐々に拡大する映像と刺激間時間間隔 (ISI) や持続時間を操作した音とは調和しない。

Telop 7 は、文字列の中心の「m」を軸に右方向 (時計回り) へ 1 回転する。Telop 7 は、Sound 6 (上昇音高) との組み合わせで調和度が高いが、他のサウンド・パターンとの組み合わせでは調和度が低い。この結果は、右方向へ円運動をする映像と連続的に音高が上昇する音が合うことを示す。一般に、音の高さは音色的高さ (tone height) と音楽的高さ (tone chroma) の二つの性質に区別される (e. g., Bachem, 1950; Révész, 1954)。Ueda and Ohgushi (1987) は多次元尺度構成法 (MDS) を用い、音高の垂直の次元 (i. e., tone height) と循環の次元 (i. e., tone chroma) の存在を実証した。円運動をする映像は、音の高さの変化と対応する可能性がある。この様に Telop 7 と Sound 6 の組み合わせから知覚された調和感は、音と映像の空間的変化パターンの類似による構造的調和に基づくものと考えられる。一方、Telop 7 は Sound 3 (断続音), Sound 4 (単音) のような刺激間時間間隔 (ISI) や持続時間が短い音とは調和しない。

Telop 8 は、実験 1 の Telop 7 を参考にして合成した。Telop 8 は、2000 ms にかけて徐々に明るくなり、輝度 (明るさ) のピーク (max) 値は 11.46 cd/m^2 である。Telop 8 は、Sound 8 (増大音), Sound 9 (増大減少音) と調和した。映像の明るさは視覚刺激の強さを反映し、音の大きさは聴覚刺激の強さを反映する (e. g., Marks, 1989; Stevens, 1957)。第 1 章で述べたように、視覚領域の明るさは聴覚領域のラウドネスと深い関係を持つ。Telop 8 と Sound 8 の組み合わせは、視覚と聴覚の強さ (エネルギー) の変化

パターンが類似している。この視聴覚組み合わせから知覚された調和感は、強さの側面における音と映像の変化パターンの類似による構造的調和に基づくものと考えられる。Marks (1989) は、変化のない光と音を用い、視覚の明るさと音の高さと大きさとの対応関係を調べた。その結果、音の高さと大きさの両方が「様相間類似性 (cross-modal analogues)」として視覚の明るさと対応したが、音の大きさより高さの方が顕著であった。しかし、本実験では、視覚の明るさと音の高さの対応関係は明確ではない。Telop 8 と音高が連続的に上昇する Sound 6 との組み合わせの平均調和評定値は 3 (どちらでもない) であるが、Telop 8 と音の大きさが徐々に増大する Sound 8 との組み合わせの平均調和評定値は高い。この様に視覚の明るさが変化する場合は、音の大きさと対応する傾向がある。一方、Sound 9 の場合は 1000 ms を境に振幅エンベロープが徐々に増大してから減少しており、映像の明るさと音のラウドネスの変化の後半部が合わなかったにもかかわらず調和度が高い。人は、視覚と聴覚のエネルギーの増加部が対応することで、調和を感じるかもしれない。このことは、音と映像の強さの変化の類似における構造的調和の場合、両者の前半部の変化が重要であることを示唆する。しかし、Telop 8 と Sound 5 (単音列) の組み合わせは、最も調和しない。徐々に明るくなるテロップ・パターンと単音が連続的に流れるサウンド・パターンは合わない。さらに、Telop 8 は、Sound 3 (断続音)、Sound 4 (単音)、Sound 7 (FM 音)、Sound 10 (random 音高) のような刺激間時間間隔 (ISI) や持続時間、周波数を操作した音とも調和度が低い。

Telop 9 は、実験 1 の Telop 2 を参考に合成した。Telop 9 は、1 回ゆっくりと光ってから消えるパターンであり、輝度のピーク値は 26.05 cd/m^2 である。Telop 9 は Sound 9 (増大減少音) との組み合わせで調和度が高い。明るさが 1000 ms を境に徐々に増大してから減少するテロップ・パターンと振幅エンベロープが 1000 ms を境に徐々に増大してから減少するサウンド・パターンが合っている。この視聴覚組み合わせから知覚された調和感は、本実験の Telop 8 と Sound 8 の組み合わせと同様に音と映像の強さの変化パターンの類似による構造的調和に基づくものである。さらに、Telop 9 は Sound 8 (増大音) との組み合わせでも調和度が高いが、映像の明るさと音の大きさの後半部の変化が合っていない。このことは、本実験の Telop 8 と Sound 9 の組み合わせと同様に、音と映像の強さの変化の類似における構造的調和は、両者の前半部の変化が重要であることを表す。一方、Telop 9 と Sound 2 (点滅音)、Sound 3 (断続音)、Sound 4 (単音)、

Sound 5 (単音列), Sound 7 (FM 音), Sound 10 (random 音高) との組み合わせの調和度は低い。この様に Telop 9 は, 刺激間時間間隔 (ISI) や持続時間, 周波数を操作した音とは調和しない。

Telop 10 は, 7 個の文字が複雑に変化する。Telop 10 は Sound 10 (random 音高) との組み合わせで調和度が高い。Sound 10 は, 持続時間 100 ms の 20 個の半音階 (B4 から F#6) をランダムに分布させ, 調性感や旋律感が生じないようにした。この視聴覚組み合わせは, 音と映像の複雑な変化パターンの類似による構造的調和が要因で, 高い調和度が得られたと考えられる。しかし, Telop 10 は Sound 2 (点滅音), Sound 3 (断続音), Sound 4 (単音) といった刺激間時間間隔 (ISI) や持続時間を操作した音との組み合わせでは調和度が低い。

以上の Telop 5, Telop 6, Telop 7, Telop 8, Telop 9, Telop 10 と調和するサウンド・パターンは, 各映像刺激の空間的/強さ (明るさ) の変化パターンと類似することが分かった。なお, 音と映像の強さの変化の「前半部」が合う場合は, その「後半部」が合わなくても高い調和度を得られた。一方, 音と映像の変化パターンが類似していない場合は, 被験者の調和評定値は低くなった。例えば, Sound 7 (FM 音) の場合, すべての映像刺激との組み合わせで調和度が低かった。また, Sound 2 (点滅音), Sound 3 (断続音), Sound 4 (単音) といった刺激間時間間隔 (ISI) や持続時間を操作した音とも調和しなかった。これらの結果は, 音と映像の変化パターンの類似による構造的調和が視聴覚情報の主観的調和に貢献することを示すものである。

本実験では, 合成したテロップと効果音を用いて印象評定実験を行い, 音と映像の構造的調和の要因を検討した。その結果, テロップと効果音の時間構造の同期と変化パターンの一致に基づく構造的調和が音と映像の主観的調和を高めた。しかし, 2.3.3.1 項で検討したように, 本実験で操作したテロップ・パターンとサウンド・パターンは特定の意味的側面 (印象) を持っている。そのため, 本研究で得られた視聴覚刺激の主観的調和には, 音と映像の印象の類似による意味的調和の影響もあった可能性がある。そこで, 意味的調和の影響があったかどうかを検討するため, 次式から得られた音と映像の印象の差の絶対値 (0 に近いほど同じ印象) と視聴覚刺激の平均調和評定値との相関係数を求めた。その結果を表 2.7 に示す。

$$| V_{(impression)} - (A_{(impression)}) | \dots\dots\dots (2.1)$$

$V_{(impression)}$: 各映像刺激の印象に対する主成分得点

$A_{(impression)}$: 各音刺激の印象に対する主成分得点

表 2.7 より、視聴覚刺激の平均調和評定値とその映像刺激と音刺激の印象の差の絶対値の相関係数は、「力動」「評価」において統計的に有意ではない。一方、「活動」においては統計的に有意な相関がみられたが、その値はかなり低い。従って、本実験で示された音と映像の主観的調和には「力動」「活動」「評価」などの印象の類似によってもたらされる意味的調和の影響は僅かであったといえる。

以上のように、実験 2 では 2 種類の音と映像の構造的調和を明らかにした。一つは、音と映像の時間的側面における時間構造の類似である。もう一つは、音と映像の空間的/強さ（明るさ）の側面における変化パターンの類似である。音と映像の時間構造の類似に基づく構造的調和は、一つの映像に一つの音が調和する傾向がある。例えば、Telop 2, Telop 3 と調和した Sound 2, Sound 3 は、映像と類似した時間構造を有する。しかし、Telop 1 と Telop 4 の場合、一つの映像に二つの音が調和したが、これらの視聴覚刺激では音と映像の「onset」が同期していた。すなわち、音と映像の時間構造の「onset」の同期が、視覚事象と聴覚事象の時間次元における構造的調和を形成するのに強く影響することが示唆された。

一方、音と映像の変化パターンの類似に基づく構造的調和は、一つの映像に複数の音が調和する傾向がある。例えば、Telop 6 と調和した Sound 6, Sound 8 は、映像の変化

表 2.7 実験 2 の映像刺激と音刺激の各主成分に対する主成分得点の差の絶対値と平均調和評定値の相関係数

視聴覚刺激の調和	映像刺激と音刺激の主成分得点の差の絶対値		
	力動	活動	評価
平均調和評定値	- .19	-.31**	.08

** 1% 水準で有意

パターンと類似した音の高さと大きさの変化パターンを有する。これらの結果は、音と映像の変化パターンの類似の「多義性 (ambiguity)」が関係すると考えられる。Marks (1989) は、異なる感覚モダリティにおける「様相間結合 (cross-modal connection)」の属性として、「一元的連合 (unitary association: 一対一)」と「二元的連合 (dual associations: 一対多)」があると提案した。Lipscomb and Kim (2004) は、音楽パラメータと映像パラメータの間で「非一元的関係 (non-unitary relationship)」があることを主張した。Eitan and Granot (2004, 2006) も、各音楽パラメータがいくつかの運動パラメータと結び付けられると主張した (one-to-many mapping)。従って、本研究においてもこれらの研究と同様の傾向がみられたものと考えられる。

音と映像の時間構造は、時間軸上のある一点と一点を捉えたものであり、音と映像の変化パターンは、変化に伴う時間より変化の形を捉えたものと考えられる。この様に、音と映像の時間構造の類似と変化パターンの類似に基づく構造的調和は、異なる要因によって主観的調和を形成することが分かる。

2.4 実験3: 2.3.3.3項で示した構造的調和の要因に関する追加実験

2.4.1 はじめに

2.3節の実験2では、テロップと効果音の時間構造と変化パターンの類似に基づく構造的調和が音と映像の主観的調和を上昇させ、視聴覚情報の評価を高めることを示した。

しかし、音と映像の構造的調和の要因について、いくつか不明な点もあった。例えば、

(i) テロップと効果音の時間構造/変化パターンの類似に基づく構造的調和の場合、音と映像の開始部/前半部と終止部/後半部のどちらが調和感に大きく貢献するのか。

(ii) テロップと効果音の空間的変化パターンの類似に基づく構造的調和の場合、その調和感に「対称性 (or 非対称性)」は存在するのか。もしそうであれば、拡大する映像には音の高さ/大きさの増加が調和し、縮小する映像には音の高さ/大きさの減少が調和することになる。一方、その逆の場合 (e. g., 映像の拡大と音の高さ/大きさの減少) は、調和しないことになる。

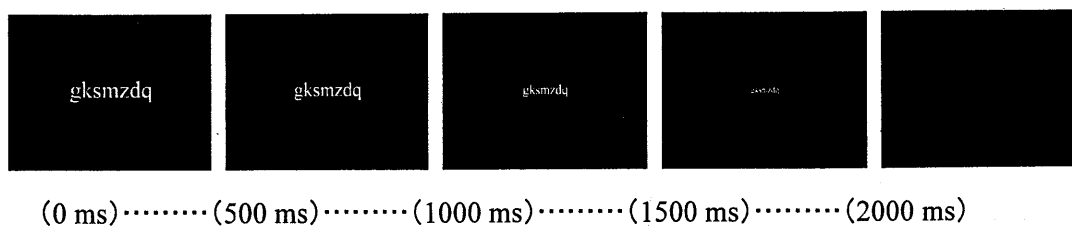
(iii) テロップと効果音の強さの変化パターンの類似に基づく構造的調和の場合、その調和感に対称性 (or 非対称性) は存在するのか。つまり、明るくなる/暗くなる映像に音の大きさの増大/減少が調和し、明るくなる/暗くなる映像に音の大きさの減衰/増大は調和しないのか。(iv) 右方向へ1回転(時計回り)する映像の変化は音の高さ(上昇と下降)、FM音の周期(1回から3回)、オクターブ循環性

(1 octave から 3 octave) を操作した音の変化と調和するのか。以上の疑問に答えるためには、対称となる視聴覚刺激が必要となる。本節では、系統的に操作した実験刺激を作成して追加実験を行い、音と映像の構造的調和の要因について考察を加える。

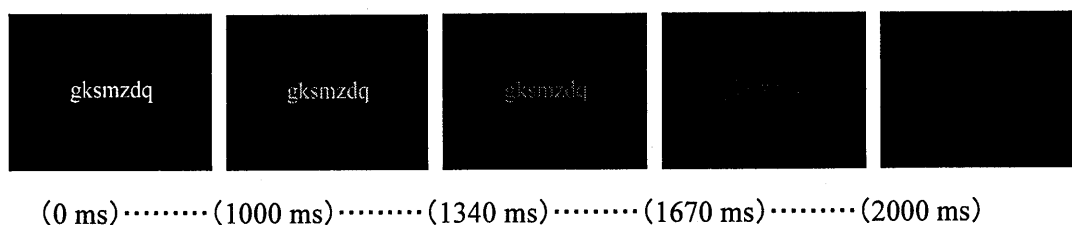
2.4.2 実験方法

被験者は、実験1と実験2に参加しなかった九州大学の音響設計学科の学部生および大学院生10名である。年齢は21歳から34歳(平均年齢23.7歳)であり、男性が8名、女性が2名である。すべての被験者は正常な視力(矯正視を含む)と聴力を有しており、映像に関する特別な訓練を受けていない。

映像刺激の作成は、Adobe社のPhotoshop 7.01とAfter Effects 5.0を用いた。映像刺激は、実験2のTelop 1 (still), Telop 6 (zoom in), Telop 7 (rotation), Telop



a. Telop 6' : Zoom out



b. Telop 8' : Fade out

図 2.18 実験 3 で新たに作成した映像刺激とその特徴

8 (fade in), Telop 9 (fade in & out) に, 図 2.18 に示すように Telop 6 の比較対照条件として徐々に小さくなる Telop 6' (zoom out) と Telop 8 の比較対照条件として徐々に暗くなる Telop 8' (fade out) を追加した。追加した映像刺激の持続時間も, 他の映像刺激と同様に 2000 ms である。

音刺激の作成は, Cybernet Systems 社の MATLAB 6.0 を用いた。音刺激は, 実験 2 の Sound 1 (定常音), Sound 4 (単音), Sound 6 (上昇音高), Sound 7 (FM 音), Sound 8 (増大音), Sound 9 (増大減少音) に, Sound 6 の比較対照条件として音の高さが連続的に下降する Sound 6' (下降音高) と Sound 8 の比較対照条件として音の大きさが徐々に減少する Sound 8' (減少音) を追加した。また, Sound 7 は, 2 周期の変調を持っていたが, 新たに 1 周期と 3 周期の変調を持つ FM 音として Sound 7' と Sound 7'' を合成した。基本周波数 (1 kHz) と変調度 (100) は Sound 7 と同じであるが, 0.5 Hz と 1.5 Hz のようにそれぞれ変調周波数のみが異なる。さらに, 音高が 1 オクターブ (C4 から C5), 2 オクターブ (C4 から C6), 3 オクターブ (C4 から C7) 変化するグライド音も追加し, 各々に Sound 11, Sound 11', Sound 11'' と名づけた。Sound 4 (200 ms) を除き, 各音刺激の持続時間は 2000 ms である。すべての音刺激には, クリック音の発生を防ぐために 5 ms の直線的な立ち上がり, および減衰特性をつけている。追加した

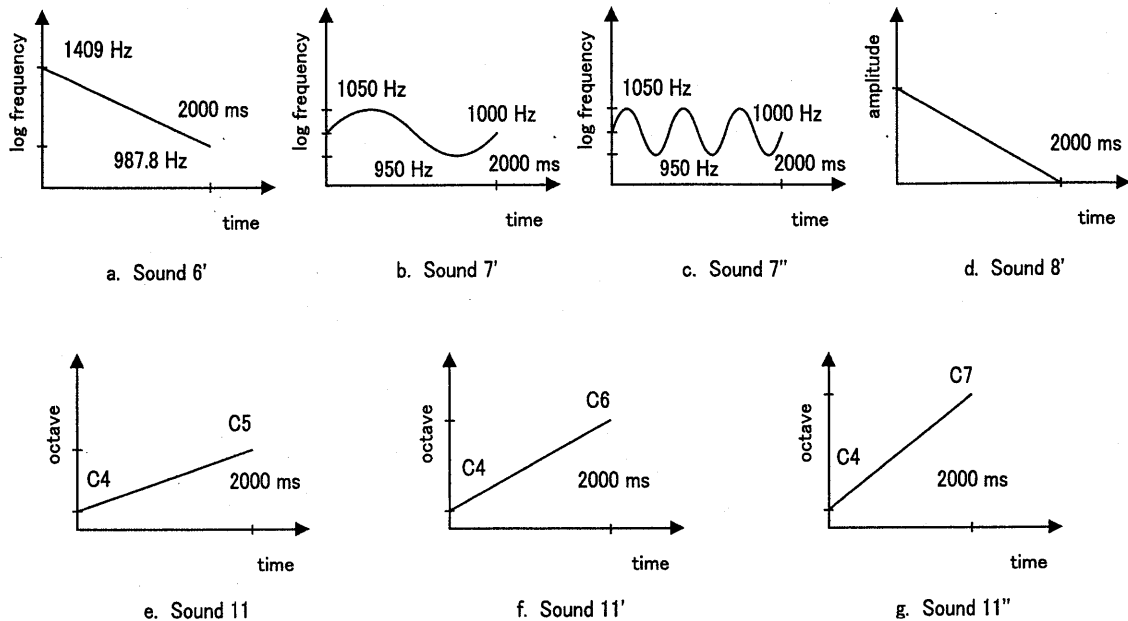


図 2.19 実験 3 で新たに作成した音刺激とその特徴

音刺激の特徴を図 2.19 に示す。

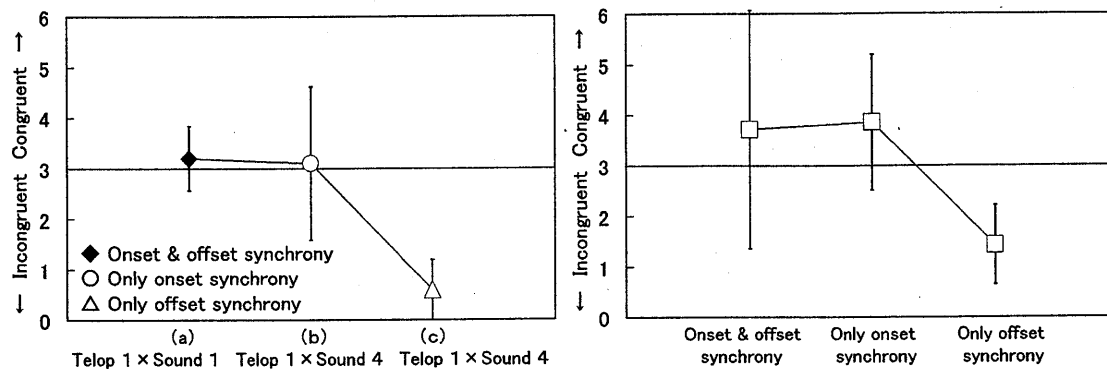
Sound 1, Sound 4, Sound 6, Sound 7, Sound 8, Sound 9 の各音刺激の呈示音圧レベル (等価騒音レベル) は, 実験 2 と同じである。追加した各音刺激の呈示音圧レベルは, Sound 6', 83 dB; Sound 7', 84 dB; Sound 7'', 84 dB; Sound 8', 79 dB; Sound 11, 80 dB; Sound 11', 86 dB; Sound 11'', 85 dB であった。

映像刺激は, パーソナルコンピュータ (IBM Think Centre A50P) を用い, 19 inch のディスプレイ (IBM-6651U3N) の画面上に呈示した。音刺激は, オーディオ・インターフェイス (EDIROL AUDIO Capture UA-5) とヘッドフォンセット (STAX Lambda Nova BASIC System) を介して, 被験者の両耳に呈示した。被験者には, すべての視聴覚刺激の調和度について 7 段階の「調和していない—調和している」の尺度を用いて評定させた。

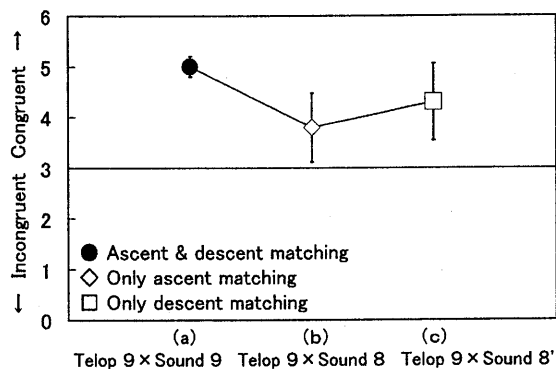
実験環境, 刺激の呈示や評定方法は, 実験 2 と同様である。

2.4.3 結果と考察

(i) の時間構造の「onset (開始)」と「offset (終止)」の効果を検討する視聴覚刺激は, Telop 1 と Sound 1 (定常音), Sound 4 (単音) を用い, (a) 音と映像の「onset」と「offset」が同期する条件, (b) 音と映像の「onset」は同期するが「offset」が同



(a) 音と映像の時間構造 (1800 ms のずれ) (b) 音と映像の時間構造 (600 ms のずれ)



(c) 音と映像の変化パターン

図 2.20 実験 3 の (i) の音と映像の時間構造の同期と変化パターンの一致の各条件での平均調和評定値の差 (エラーバーは標準偏差を表す)

期しない条件, (c) 音と映像の「offset」は同期するが「onset」が同期しない条件を設けた。(b) と (c) の非同期条件の場合, 音と映像の間に 1800 ms の時間ずれがある。この値は, 視覚事象と聴覚事象の同期ずれの検知閾と主観的許容度を十分上回るものである (第 3 章の図 3.6 参照)。印象評定実験より得られた平均調和評定値を図 2.20-a に示す。

図 2.20-a より, 相対的に (a) の音と映像の「onset」と「offset」が同期する条件と (b) の音と映像の「onset」のみが同期する条件の調和度は同程度に高い傾向がある。一方, (c) の音と映像の「offset」のみが同期する条件の場合, 非常に調和度が低い傾向がある。(a), (b), (c) の三つの条件間の調和度の相違を検討するため, 反復測定による一元配置の分散分析を行った。その結果, 視聴覚刺激間の調和感に統計的に有意な

差が認められた [$F_{(2,18)} = 12.77, p < .01$]。本実験の結果と実験1のTelop 3, 実験2のTelop 1, Telop 4の結果を合わせると、音と映像の時間構造の類似に基づく構造的調和は、両者の「終止部」より「開始部」の同期による効果が顕著であることが明確になった。また、多重比較 (Tukey's HSD 法) の結果も (c) と (a), (b) の間のみに統計的に有意な差が認められ ($p < .01$)、以上の結果を強く裏付けるものとなった。しかし、Fraisse (1982/1987) は聴覚事象間の時間的な群化の上限を 1500 ms から 2000 ms の間 (Bolton, 1894; MacDougall, 1903) であると特定しており、本実験で用いた音と映像は時間的につながっていないものとして知覚された可能性もある。そこで、Fraisse が知覚的な群化 (grouping) と精度 (precision) の最適値として挙げた 600 ms の時間ずれを用い、図 2.20-a の 1800 ms の条件と同様の実験を行った。実験には、1800 ms の条件に参加した被験者のうち、7名 (男性6名, 女性1名) が参加した。その結果を示した図 2.20-b より、図 2.20-a の結果と同様に音と映像の「onset」が同期する条件の方が「offset」が同期する条件より調和した ($p < .01$)。

一方、変化パターンの「前半部」と「後半部」の効果を検討する視聴覚刺激は、Telop 9 と Sound 9 (増大減少音), Sound 8 (増大音), Sound 8' (減少音) を用い、(a) 音と映像の変化の「前半部」と「後半部」が合う条件、(b) 音と映像の変化の「前半部」は合うが「後半部」が合わない条件、(c) 音と映像の変化の「後半部」は合うが「前半部」が合わない条件を設けた。印象評定実験より得られた平均調和評定値を図 2.20-c に示す。

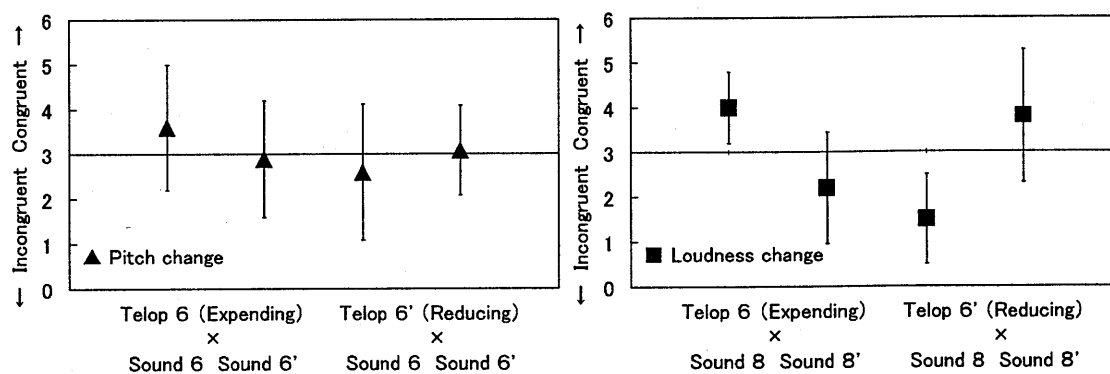
図 2.20-c より、(a), (b), (c) の三つの条件において調和度が高い傾向がある。3条件の中では、(a) の音と映像の変化の「前半部」と「後半部」が合う条件の調和度が高く、相対的に (b) の「前半部」のみが合う条件は調和度が低い傾向がある。(c) の「後半部」のみが合う条件は (a) と同程度に調和度が高い傾向がある。(a), (b), (c) の三つの条件間の調和度の相違を検討するため、反復測定による一元配置の分散分析を行った結果、視聴覚刺激間の調和感に統計的に有意な差が認められた [$F_{(2,18)} = 6.98, p < .01$]。事後比較として、Tukey (HSD) 法による多重比較を行った結果、(a) と (b) の間のみに統計的に有意な差が認められた ($p < .05$)。音と映像の変化パターンの類似による構造的調和は、両者の変化の「前半部」より「後半部」による影響が少し大きいようである。しかし、実験2のTelop 8, Telop 9の結果は、音と映像の変化の「前半

部」が合うだけで十分調和度が高かった。また、Telop 8 と Sound 9 (変化の「前半部」が一致), Telop 8' と Sound 9 (変化の「後半部」が一致) を組み合わせた視聴覚刺激を用いた実験でも、両方とも平均調和評定値が4で、視聴覚刺激の変化の「前半部」と「後半部」の相違による調和感に統計的に有意な差はなかった ($p > .1$)。

以上の結果をまとめると、テロップと効果音の時間構造の類似に基づく構造的調和は、両者の「終止部」より「開始部」の同期が重要である。しかし、難波、林、若生 (2003) が行った画面上で円盤が移動する映像と音高の変化の同期に関する実験では、両者の「始点」より「終点」の同期が優位であった。すなわち、映像が出現して単に消える場合は、その「開始部」と音の同期が重要であり、映像が移動して停止する場合は、その「終点部」と音の同期が重要であると考えられる。一方、音と映像の強さの変化パターンの類似に基づく構造的調和は、「前半部」と「後半部」の一致の相違による調和感の変化は僅かである。

(ii) の拡大する/縮小する映像と音の高さの上昇/下降、音の大きさの増大/減少の調和に関する印象評定実験には、Telop 6 と Telop 6' に Sound 6 (上昇音高), Sound 6' (下降音高), Sound 8 (増大音), Sound 8' (減少音) を組み合わせた視聴覚刺激を用いた。印象評定実験より得られた結果を図 2.21 に示す。

まず、図 2.21-a の平均調和評定値を見ると、拡大するテロップと音の高さの上昇 (Telop 6 × Sound 6), 縮小するテロップと音の高さの下降 (Telop 6' × Sound 6') のように音と映像の変化パターンが一致する視聴覚刺激同士は、相対的に調和する側に布



(a) 映像の拡大/縮小と音高の上昇/下降 (b) 映像の拡大/縮小と音量の増大/減少

図 2.21 実験 3 の (ii) の映像のサイズの変化と音の大きさと高さの変化の各条件での平均調和評定値の差 (エラーバーは標準偏差を表す)

置している。一方、音と映像の変化パターンが正反対である視聴覚刺激同士 (Telop 6 × Sound 6', Telop 6' × Sound 6) は相対的に調和していない側に布置している。しかし、全体的に各視聴覚刺激間の調和度の差はないようである。実験要因による調和度の相違を検定するため、映像 (2 水準) × 音 (2 水準) の反復測定による二元配置の分散分析を行った。その結果、映像の主効果と音の主効果は統計的に有意な差はみられず、映像刺激 × 音刺激の交互作用 [$F_{(1,9)} = 5.09, p < .05$] のみが統計的に有意であった。Tukey (HSD) 法による多重比較を行い、視聴覚刺激の水準間の調和感の相違を検討したが、統計的に有意な差はみられなかった ($p > .1$)。

次に、図 2.21-b より、拡大するテロップと音の大きさの増加 (Telop 6 × Sound 8), 縮小するテロップと音の大きさの減少 (Telop 6' × Sound 8') のように音と映像の変化パターンが一致する視聴覚刺激同士は調和する側に布置しているが、音と映像の変化パターンが正反対である視聴覚刺激同士 (Telop 6 × Sound 8', Telop 6' × Sound 8) は調和していない側に布置している。映像の拡大/縮小と音の大きさの増大/減少を要因とした二元配置の分散分析を行った結果、映像の主効果と音の主効果は統計的に有意な差はみられず、映像刺激 × 音刺激の交互作用 [$F_{(1,9)} = 10.74, p < .01$] のみが統計的に有意であった。各視聴覚刺激の調和度の相違を検討するため、Tukey (HSD) 法による

表 2.8 実験 3 の (ii) の拡大/縮小する映像と増大/減少する音を組み合わせた視聴覚刺激の調和感に関する多重比較の結果

視聴覚刺激	視聴覚刺激		
	調和		
調和	Telop 6 × Sound 8'	Telop 6' × Sound 8	Telop 6' × Sound 8'
Telop 6 × Sound 8	1.80*	2.50**	n. s.
Telop 6 × Sound 8'	-	n. s.	-1.60 [^]
Telop 6' × Sound 8	n. s.	-	-2.30**
Telop 6' × Sound 8'	1.60 [^]	2.30**	-

[^]10% *5% **1% 水準で有意

多重比較を行った。表 2.8 の多重比較の結果より、音と映像の変化パターンが一致する視聴覚刺激同士と正反対である視聴覚刺激の間の調和感に統計的に有意な差が認められた。ただし、Telop 6 と Sound 8' の変化パターンが正反対である視聴覚刺激と Telop 6' と Sound 8' の変化パターンが一致する視聴覚刺激では、10%水準で有意である。一方、音と映像の変化パターンが一致する視聴覚刺激同士、音と映像の変化パターンが正反対である視聴覚刺激同士には統計的に有意な差はなかった ($p > .1$)。

以上のように、映像のサイズの拡大/縮小と音の大きさの増大/減少との対応関係は明確であり、音と映像の変化パターンにおける調和感の「対称性」が示された。つまり、拡大する映像には音の大きさの増大は調和するが、減少は調和しない。逆に、縮小する映像には音の大きさの減少が調和するが、増大は調和しない。この様にテロップと効果音の変化パターンの類似が構造的調和を形成し、音と映像の主観的調和の上昇に貢献することが確認できた。

しかし、実験 2 の図 2.17-f では拡大する映像 (Telop 6) に音高の上昇 (Sound 6) が調和したが、本実験の図 2.21-a では映像のサイズの拡大/縮小と音の高さの上昇/下降との対応関係は明確ではなかった。一方、実験刺激は異なるが、有田、岩宮 (2004)、岩宮ら (2003, 2005) の研究では、拡大する映像と上昇形の音高、縮小する映像と下降形の音高が調和する結果が得られた。本実験では、各視聴覚刺激に対する平均調和評定値は「どちらでもない (3)」の付近に布置しており、本実験の被験者たちは音の高さの変化は映像の垂直の変化と最も対応するか、映像の奥行きの変化は音の大きさの変化と最も対応するというような絶対的な判断基準を持っていたかもしれない。

(iii) の明るくなる/暗くなる映像と音の大きさの増大/減少の調和に関する印象評定実験には、Telop 8 と Telop 8' に Sound 8 (増大音)、Sound 8' (減少音) を組み合わせた視聴覚刺激を用いた。印象評定実験より得られた平均調和評定値を図 2.22 に示す。

図 2.22 より、明るくなる映像と音の大きさの増大 (Telop 8 × Sound 8) と明るくなる映像と音の大きさの減少 (Telop 8 × Sound 8') の間の調和度は差がなく、暗くなる映像と音の大きさの増大 (Telop 8' × Sound 8) と暗くなる映像と音の大きさの減少 (Telop 8' × Sound 8') の間の調和度には差がある傾向がみられる。実験要因による調和度の相違を検定するため、反復測定による二元配置の分散分析を行った。その結果、映像刺激の主効果 [$F_{(1,9)} = 6.93, p < .05$], 音刺激の主効果 [$F_{(1,9)} = 10.71, p < .01$],

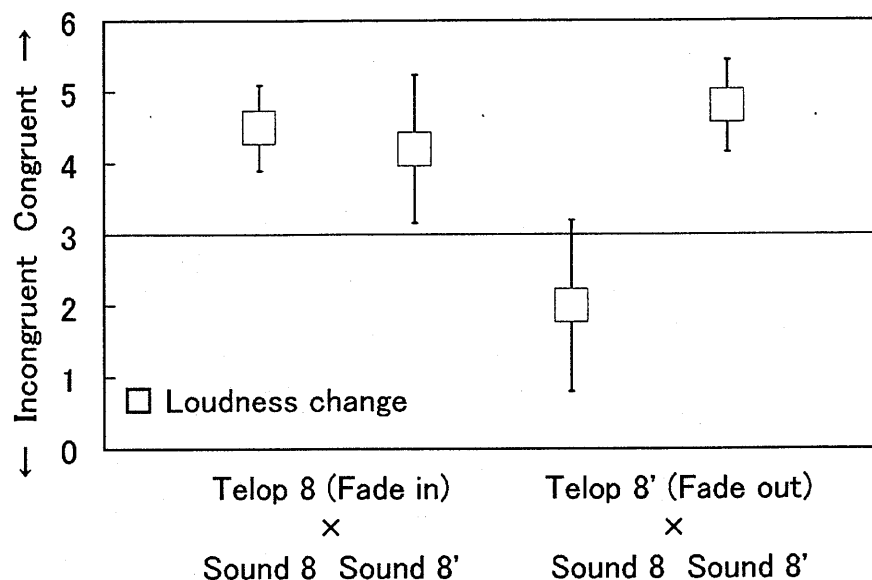


図 2.22 実験 3 の (iii) の映像の明るさの変化 (明暗) と音の大きさの変化 (大小) の各条件での平均調和評定値の差 (エラーバーは標準偏差を表す)

映像刺激×音刺激の交互作用 [$F_{(1,9)} = 12.93, p < .01$] が, それぞれ統計的に有意であった。各視聴覚刺激の調和度の相違を検討するため, Tukey (HSD) 法による多重比較を行った結果, 最も調和度の低い Telop 8' と Sound 8 の組み合わせと他の三つの組み合わせ (Telop 8×Sound 8, Telop 8×Sound 8', Telop 8' ×Sound 8') との間だけに統計的に有意な差が認められた ($p < .01$)。すなわち, 音と映像の変化パターンが合う視聴覚刺激同士 (Telop 8×Sound 8, Telop 8' ×Sound 8') は同程度に調和度が高い。明るくなる映像には大きくなる音が, 暗くなる映像には小さくなる音が調和し, 音と映像の強さの変化パターンの類似に基づく構造的調和が確認できた。一方, 音と映像の変化パターンが合わない視聴覚刺激同士の調和度は統計的に有意な差があり, 暗くなる映像と大きくなる音 (Telop 8' ×Sound 8) は調和しないと評定されたが, 明るくなる映像と小さくなる音 (Telop 8×Sound 8') は音と映像の変化パターンが一致する視聴覚刺激 (Telop 8×Sound 8, Telop 8' ×Sound 8') と同程度に調和した。この結果は, 音と映像の強さの変化の類似に基づく構造的調和に「非対称性」が存在することを示唆する。

(iv) の右方向へ 1 回転 (時計回り) する映像と音の高さの上昇/下降の調和に関する

印象評定実験には、Telop 7 に Sound 6（上昇音高）と Sound 6'（下降音高）を組み合わせた視聴覚刺激を用いた。実験 2 の図 2.17-g では時計回りの映像（Telop 7）に音高の上昇（Sound 6）が調和したが、本実験では映像の回転方向と音高の上昇/下降との対応関係はみられなかった ($p > .1$)。なお、それぞれの視聴覚刺激の平均調和評定値は 2.8 ($SD = 1.4$) と 3.2 ($SD = 1.2$) である。有田、岩宮 (2004) の上昇形/下降形の音列と場面が時計回り/反時計回りに切り替わるパターンを用いた印象評定実験でも、すべての視聴覚刺激の調和度がかなり高く、円運動をする映像の変化方向と音高の変化との対応関係は明確ではなかった。さらに、有田 (2006) は時計回り/反時計回りの映像刺激と音高が上昇/下降、上昇下降/下降上昇する音刺激を用い、一対比較法で音と映像の構造的調和を調べたが、被験者の調和判断が二つに分かれてしまい、明快な結論が得られなかった。

また、1 回転（時計回り）する映像の周期と音の周期的変化との調和を検討するため、Telop 7（1 周期）と Sound 7'（1 周期）、Sound 7（2 周期）、Sound 7''（3 周期）を用い、(a) 音と映像の変化周期が合う条件、(b) 音と映像の変化周期が 1 回ずれる条件、(c) 音と映像の変化周期が 2 回ずれる条件を設けた。この様に (a) から (c) へゆくほど、音刺激と映像刺激の変化周期のずれが大きくなる。図 2.23-a に、印象評定実験より得られた平均調和評定値を示す。

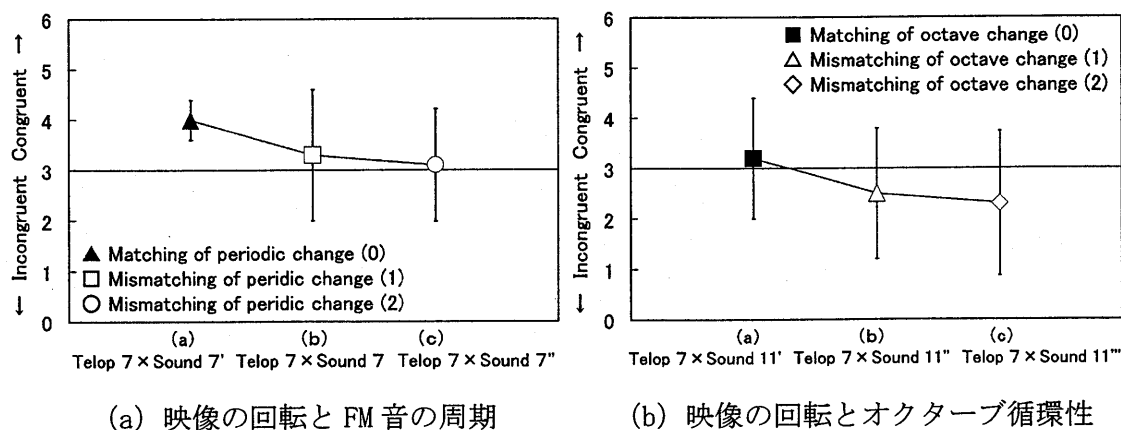


図 2.23 実験 3 の (iv) の右方向へ回転する映像の変化と音の高さの変化の各条件での平均調和評定値の差（エラーバーは標準偏差を表す）

図 2.23-a より、(a) の音と映像の変化周期が合う視聴覚刺激では調和度が高く、(b) から (c) へと音と映像の変化周期のずれが大きくなるにつれて調和度が低下する傾向がある。(a)、(b)、(c) の三つの条件間の調和度の相違を検討するため、反復測定による一元配置の分散分析を行った。その結果、(a)、(b)、(c) の間の調和感に統計的に有意な差はみられなかった ($p > .1$)。

さらに、1 回転 (時計回り) する映像の循環と音階のオクターブ循環性 (i. e., tone chroma) との調和を検討するため、Telop 7 (1 循環) と Sound 11 (1 循環), Sound 11' (2 循環), Sound 11'' (3 循環) を用い、(a) 音と映像の循環の変化が合う条件、(b) 音と映像の循環の変化が 1 回ずれる条件、(c) 音と映像の循環の変化が 2 回ずれる条件を設けた。この様に (a) から (c) へゆくほど、音刺激と映像刺激の循環の変化ずれが大きくなる。図 2.23-b に、印象評定実験より得られた平均調和評定値を示す。

図 2.23-b より、相対的に (a) の音と映像の循環の変化が合う視聴覚刺激では調和度が高く、(b) から (c) へと音と映像の循環の変化ずれが大きくなるにつれて調和度が低下する傾向がある。(a)、(b)、(c) の三つの条件間の調和度の相違を検討するため、反復測定による一元配置の分散分析を行った結果、視聴覚刺激間の調和感に統計的に有意な差が認められた [$F_{(2,18)} = 4.22, p < .05$]。しかし、Tukey (HSD) 法による多重比較を行い、(a)、(b)、(c) の水準間の調和度の相違を検討したが、統計的に有意な差はみられなかった ($p > .1$)。

以上のように (iv) では、右方向へ 1 回転 (時計回り) する映像と音の高さの諸変化との構造的調和について検討した。その結果、いくつかの意味のある傾向はみられたが、明確な結論を導けなかった。なお、(ii) の拡大/縮小する映像と音の高さの上昇/下降との調和も明確な結論を導けなかった。

2.5 本章の結論

本章では、実際のテレビ番組のテロップと効果音、合成したテロップと効果音を用いて印象評定実験を行い、音と映像の時間構造と空間的/強さ（明るさ）の変化パターンの類似がもたらす構造的調和の効果を検討した。テロップと効果音の時間構造が同期したり変化パターンが一致したりする視聴覚刺激では主観的調和度が高く、構造的調和の効果が認められた。特に、時間構造の類似に基づく構造的調和の場合、その終止部（offset）より開始部（onset）の同期が音と映像の主観的調和に及ぼす影響が大きい。本研究により、音と映像の構造的調和は時間次元のみならず、特定の変化の側面とも深い関連を持つことが明らかになった。音と映像の時間構造の類似に基づく構造的調和は、一つの映像に一つの音が調和する「一対一的な対応関係」がみられる。一方、音と映像の変化パターンの類似に基づく構造的調和は、一つの映像に複数の音が調和する「一対多的な対応関係」がみられる。なお、音と映像の変化パターンの対応関係には「対称性」や「非対称性」が存在する。また、本研究では視聴覚情報の評価、力動性、活動性の評価における Marshall and Cohen の「調和-連合モデル（Congruence-Associationist Model, 1988）」を実証した。視聴覚情報の力動性は聴覚情報より、活動性は視覚情報より直接に影響を受けるが、評価に関しては音と映像の組み合わせの主観的調和によって決まる。