

データから読み解く テレビドキュメンタリー研究

2018年3月9日 NHK放送文化研究所 宮田章

パート1

今日の発表の概要

研究の概要:

個々のテレビドキュメンタリー(TD)
のテキスト内容を数量的なデータ
を用いて読み解くこと※。

データはデータでも・・・

① **テキスト内容についてのデータ**であって、視聴率など、視聴者の受け入れ方についてのデータではない。

② テキスト内容を数量的なデータに変換して読み解くことにおいて内容分析であるが、**すべてのTDテキストについて適用できる**という点で他の内容分析と異なっている※。

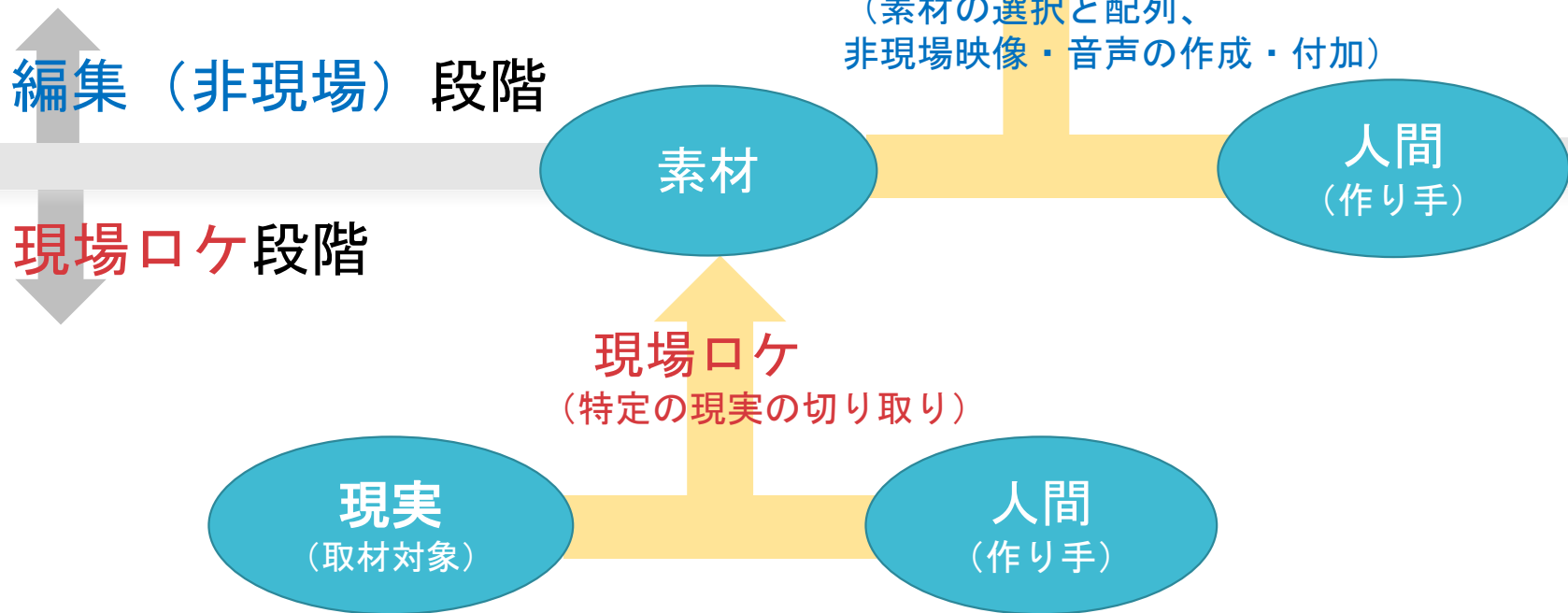
具体的には・・・

テキスト内の映像と音声を現場/非現場に、音声を言語/非言語に切り分け
それぞれの時間的分量（尺数）をカウントしてデータ化する

		音声	
		非言語音声	言語音声
現場的 ↑	現場映像	ロケ現場の物音	ロケ現場の人の声
	非現場的 ↓	非現場映像	音楽 効果音

← 非言語的 言語的 →

テレビドキュメンタリーは
現場で得られた映像・音声と
非現場で作成・付加された映像・音声との混合物である※。



データからわかること

- 1、テキスト内容のすべてがわかるわけではない。
- 2、取材対象である現実の質、撮影・録音のありよう、構成・編集の巧拙、テキスト全体のメッセージ性・・・こうしたことはわからない。
⇒従来通り質的な内容記述を行う。
- 3、そのTDが、どの程度口ケ現場の息吹を伝えているか、また、どの程度言語音声に頼っているかということがわかる。

個々のTDテキストの身長と体重がわかる

一人一人の人間がそれぞれ固有の身長・体重のデータを持つように、これらのデータは個々のTDテキストに固有のデータである。



様々なTD研究の基礎データとして有用

今日のワークショップ:前半

- 1、『ドキュメント72時間』と『NHKスペシャル』を例にとって、その映像と音声からデータをどう抽出するかを説明する。
- 2、そのデータから読み取れる各テキストの特徴を示す。

今日のワークショップ後半

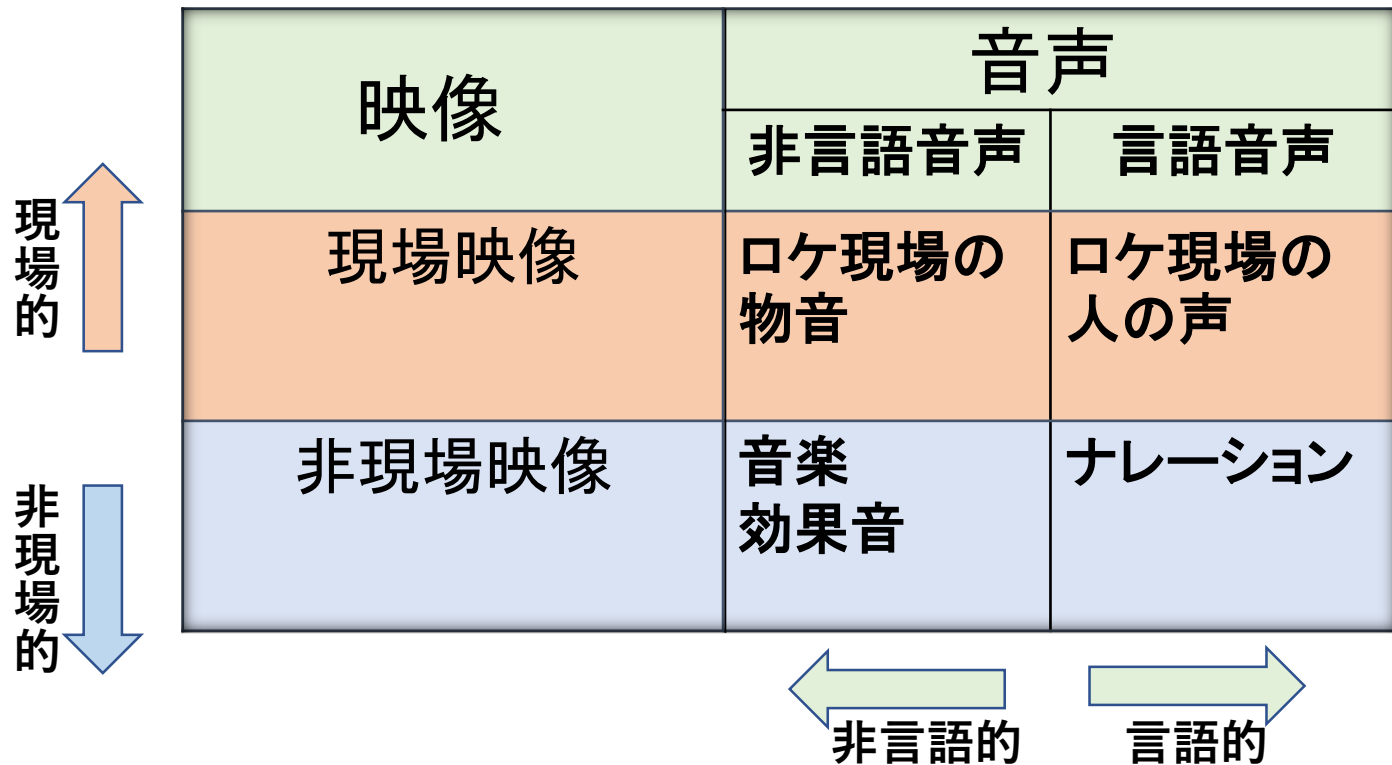
応用問題として、日本のTD史上最初のヒット作である『日本の素顔 日本人と次郎長』(1958)を視聴し、

このテキストがなぜヒットしたかを検討する。

〈パート2〉 データの抽出と分析の方法

—『ドキュメント72時間』と『NHKスペシャル』を例にして—

映像と音声を切り分ける※



映像の切り分け方(現場映像／非現場映像)

現場映像:ロケ現場の中で**具体的な場所**を占める被写体が、**通常の時間経過**の中で変化するありさまを撮影したカット。

言い換えると・・・

カメラの前に展開する現実の「その時」をカット尺の分だけ、「その場」をカメラフレームの中に入る分だけ切り取った「**その時その場**」の表象。

非現場映像

「その時その場」のダイレクトな表象ではない。

具体的なロケ現場から切り離された場所で、
作り手の操作によって生まれた映像・・・3種類

非現場映像ア

作り手が被写体をめぐる時間の流れを操作しているカット。

スチル、スロー、早送り、同じ場面の連続的なリプレー等。

非現場映像イ

作り手が被写体の場所性を抹消しているカット。
無背景のスタジオや暗幕の上など「どこでもない
場所」で撮影された映像。
代表的には文書や写真、図画などの接写。

非現場映像ウ

作り手が被写体を創造しているカット。
作り手が作成した図表などの静止画、またティラノ
サウルスの動きや、地震発生メカニズムといった
現場ロケが困難なものを可視化したCG動画。

視聴

- 1、『72時間 秋田 真冬の自販機の前で』
(2015年3月6日放送)
- 2、『Nスペ 震度7 何が生死を分けたのか』
(2016年1月17日放送)
- 3、『Nスペ 女たちの太平洋戦争 従軍看護婦
激戦地の記録』(2015年8月13日放送)

それぞれのアヴァンタイトル

3本の視聴

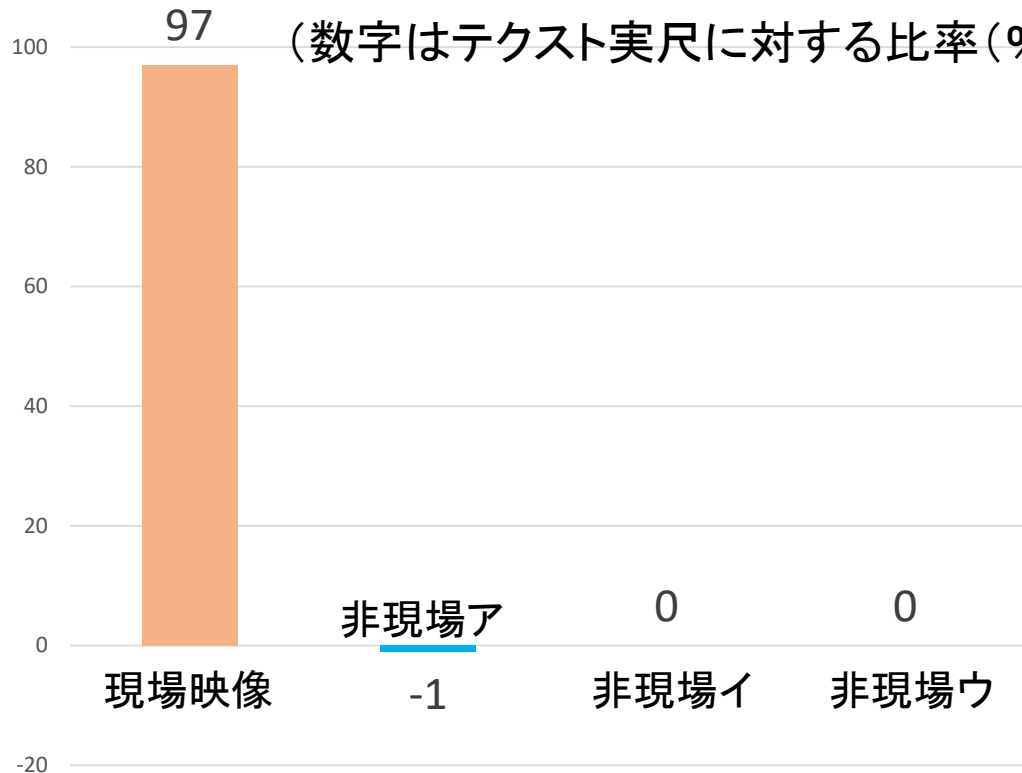
データのとりだし方

現場映像と非現場映像ア～ウのそれぞれがテキスト内に占める**時間的分量(尺数)**をカウントしてデータ化する。

各尺数が**テキスト実尺**(CMや次回予告等を除いたテキストの総尺)に占める**比率(%)**が有用。

『72時間 真冬の自販機』の映像

(数字はテキスト実尺に対する比率(%))

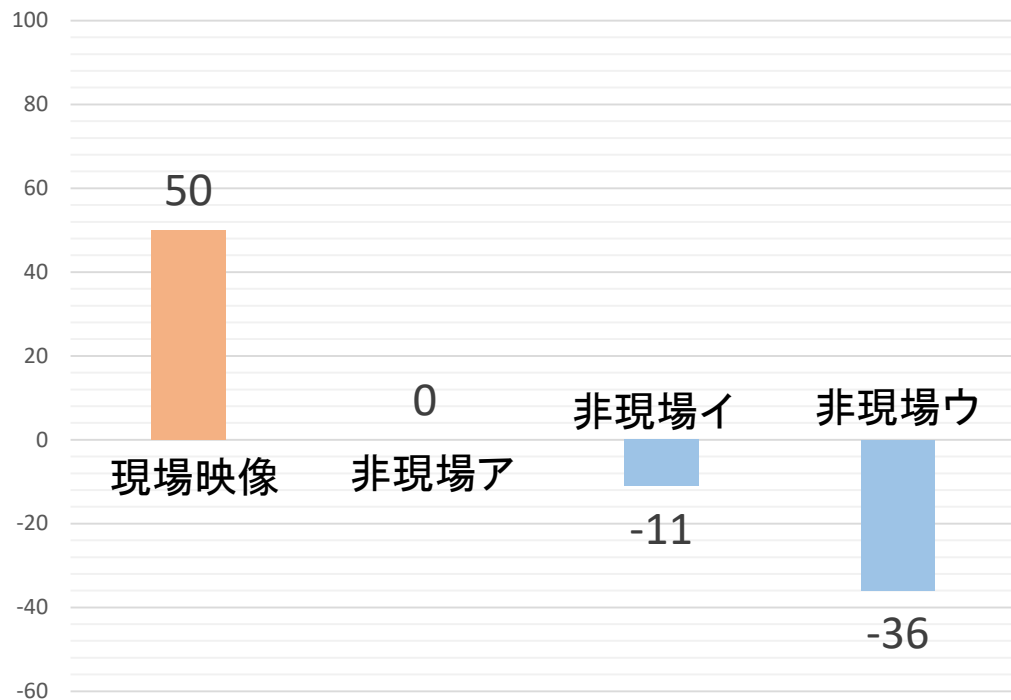


テキスト実尺1470秒のうち
現場映像1428秒、非現場映像ア30秒
イ、ウはゼロ※

顕著に**現場的**

**ロケ現場の息吹を
伝える映像**

『Nスペ 震度7』の映像

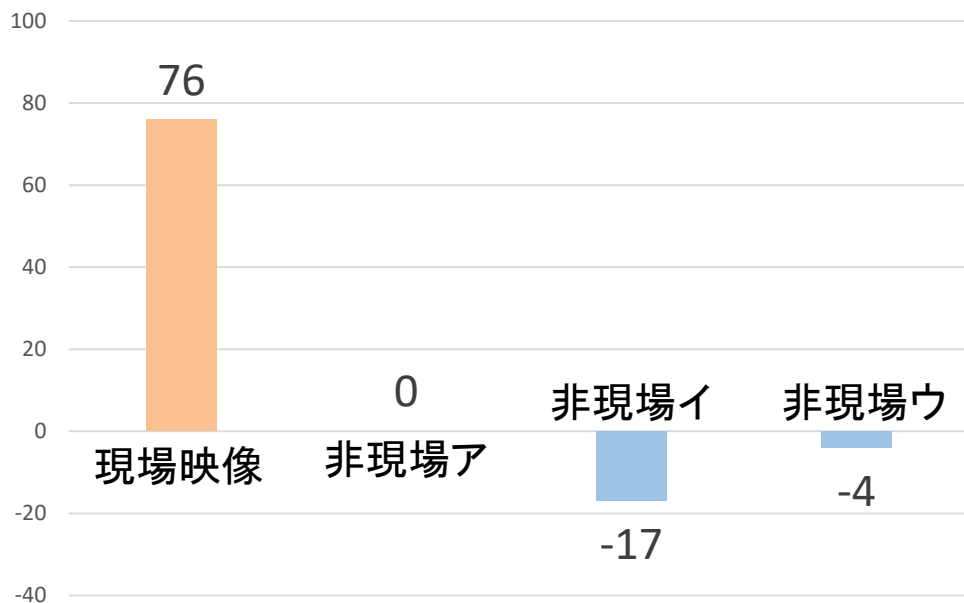


テキスト実尺2940秒のうち
現場映像1484秒、非現場ア
0秒、非現場イ335秒、非現
場ウ1047秒

TDとしては顕著に**非現場的**※

現場の息吹を伝えるというより
整理した情報をCGで伝える

『Nスペ 従軍看護婦』の映像



テキスト実尺2940秒のうち
現場映像2245秒、非現場ア
0秒、非現場イ512秒、非現
場ウ129秒

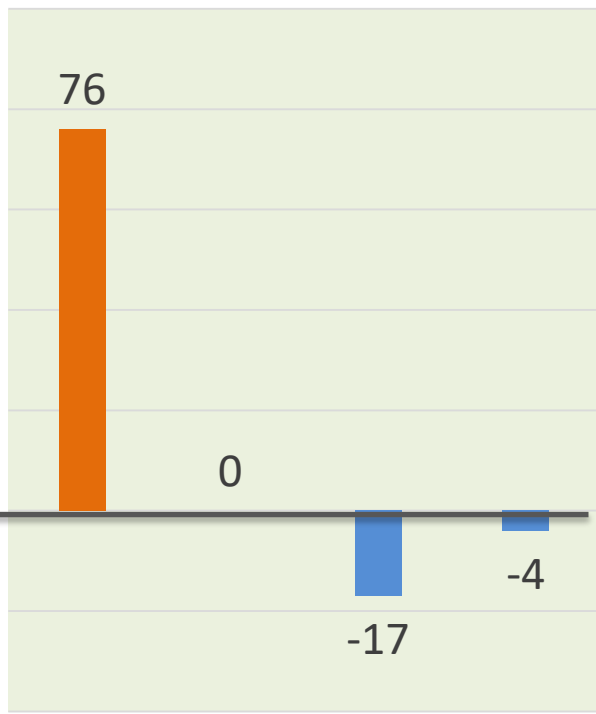
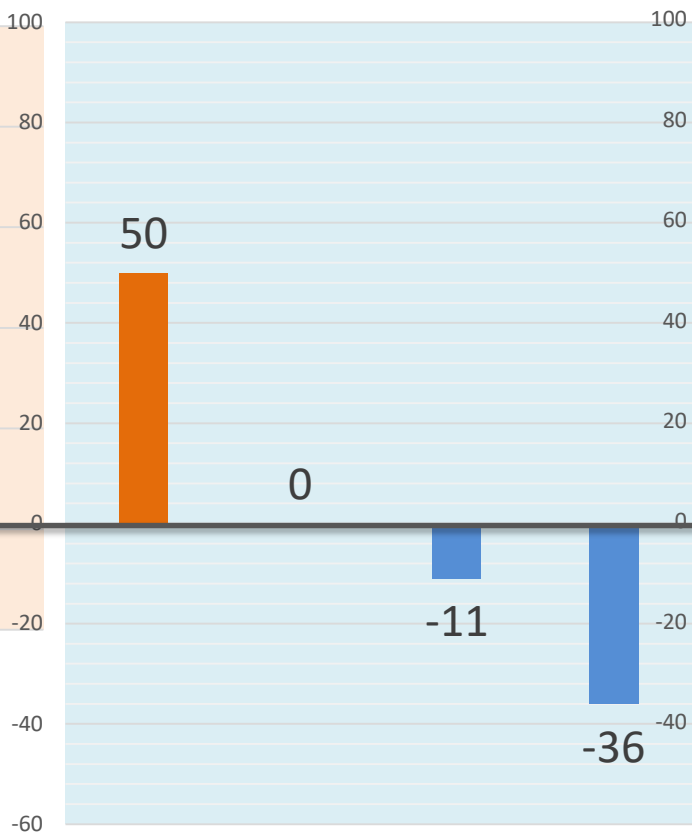
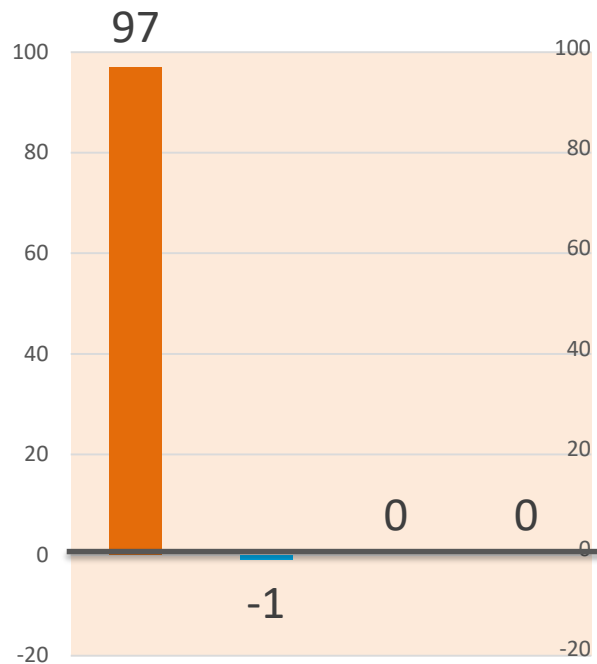
過去の検証モノ、TDの映像としては
やや非現場的。

ただし非現場映像の多くは
接写映像で、CG映像は少ない。

『72時間』

『震度7』

『従軍看護婦』



ちなみに『ドキュメント72時間』の映像データは均一性が著しい。

2015年視聴者投票ベスト5のデータ(数字は秒)

テキスト	テキスト実尺	現場映像	非現場ア	非現場イ	非現場ウ
北の大地の学生寮	1470	1442	15	0	0
赤羽おでん屋エレジー	1470	1437	21	0	0
真冬の自販機の前で	1470	1428	30	0	0
駄菓子屋	1470	1452	8	0	0
大仏を見上げる霊園で	1470	1444	12	0	0
『72時間』平均	1470 (100)	1440 (98)	17 (1)	0※	0

音声は言語音声(人の声)と非言語
音声(物音)に分けたあと、
現場／非現場に切り分ける。

音声の切り分け

- A音： **ロケ現場**で録った**物音**
(自然音、機械音、楽器音)
- B音： **ロケ現場**で録った人の**声**
(インタビューなど)
- C音： **非現場**で作成した**物音**
(音楽、効果音)
- D音： **非現場**で作成した人の**声**
(ナレーション)

音声は映像よりカウントが複雑

音量の大小があり、
A～Dの各音は1カットの中でさえしばしば重複する。

(A～D各音の尺数を単純に足し合わせると、テキスト実尺を超える)

○音量の大小は考慮せず、聞こえた音声の尺数を
すべてカウントする。

(各音を音量別にカウントするのは困難かつ煩雑※)

『72時間 真冬の自販機』の音声データ

数字はテキスト実尺に対する比率(%)

	非言語音声	言語音声
↑ 現場的	92 (A音:現場の物音)	50 (B音:現場の声)
↓ 非現場的	47 (C音:音楽など)	25 (D:ナレーション)

- 現場音(A音、B音)が顕著に多い。
ロケ現場の息吹を伝える音声構成。
- 言語音声が多い(75%)。言語的。
- A～D各音を単純に足した音声の総量はテキスト実尺の214%に達する。
(音声的に非常に“厚塗り”)
- 現場映像に多量の音楽が付加されている※(情感を「盛る」)。

『Nスぺ 震度7』の音声データ

	非言語音声	言語音声
↑ 現場的	40 (A音:現場の物音)	12 (B音:現場の声)
↓ 非現場的	46 (C音:音楽など)	42 (D:ナレーション)

○全体に**非現場音が多く**、特に言語音声で顕著である(12:42)。
ナレーションで情報を伝える音声構成。

○B音が非常に少ないために、**意外に非言語的(54%)**。

○音声の総量はテキスト実尺の140%と『72時間』に比べ**“薄塗り”**。

『Nスペ 従軍看護婦』の音声データ

	非言語音声	言語音声
現場的	50 (A音:現場の物音)	42 (B音:現場の声)
非現場的	50 (C音:音楽など)	30 (D:ナレーション)

○現場音と非現場音が拮抗している。

現場的でもあり非現場的でもある。

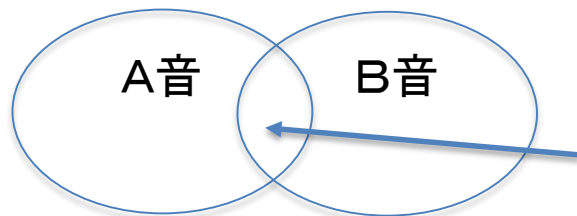
○B音が多く、D音も少なくないので、かなり言語的(72%)。

○音声の総量は172%で

『72時間』(214%)『震度7』(140%)の中間。

音声の重複分を考慮したデータ項目

現場音尺 ($A \cup B$: 現場の物音または現場の声が聞こえる尺数)
A音とB音を足し合わせたあと、両者の重複分を引いた尺数

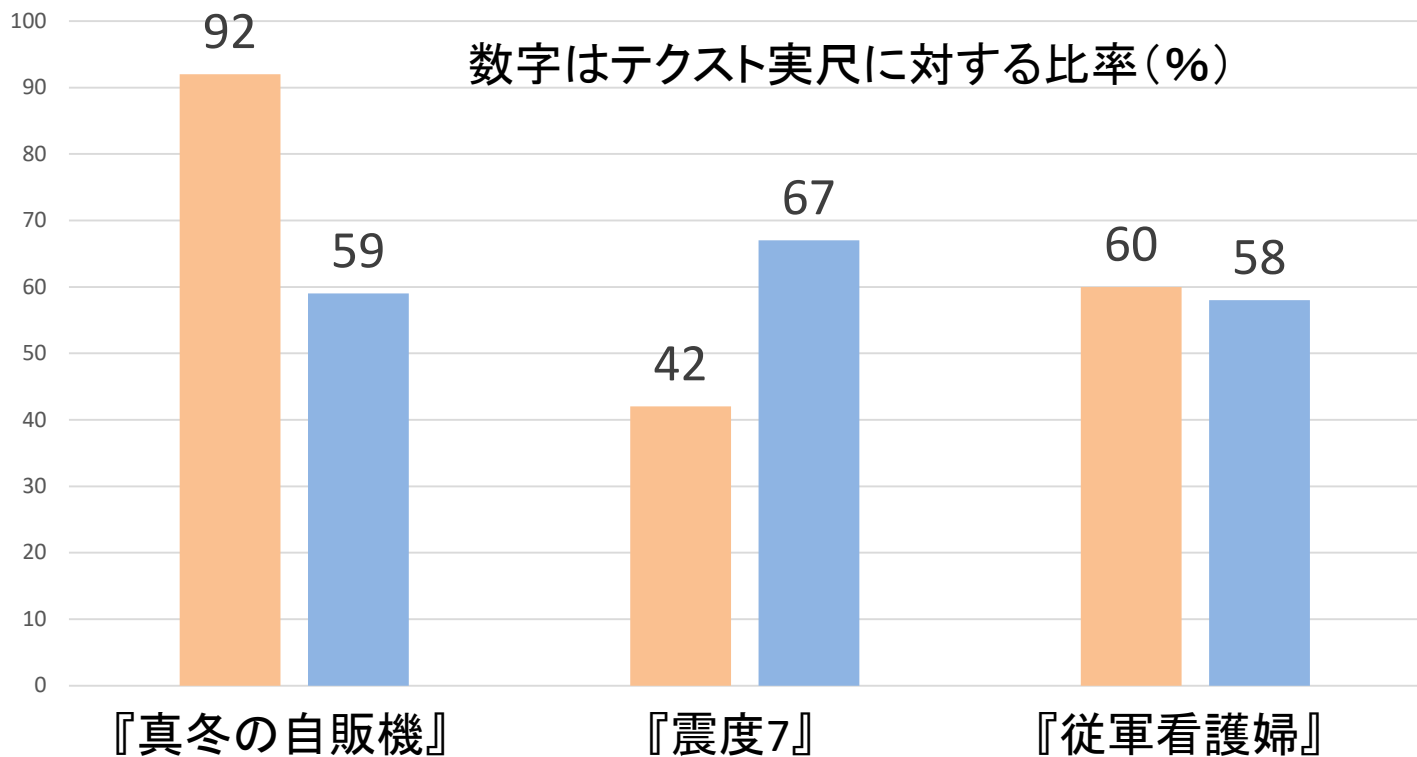


A+Bだとここを二重に数えてしまうので、一回分引く。

非現場音尺 ($C \cup D$: 音楽またはナレーションが聞こえる尺数)

言語音尺 ($B \cup D$: 現場の声またはナレーションが聞こえる尺数)

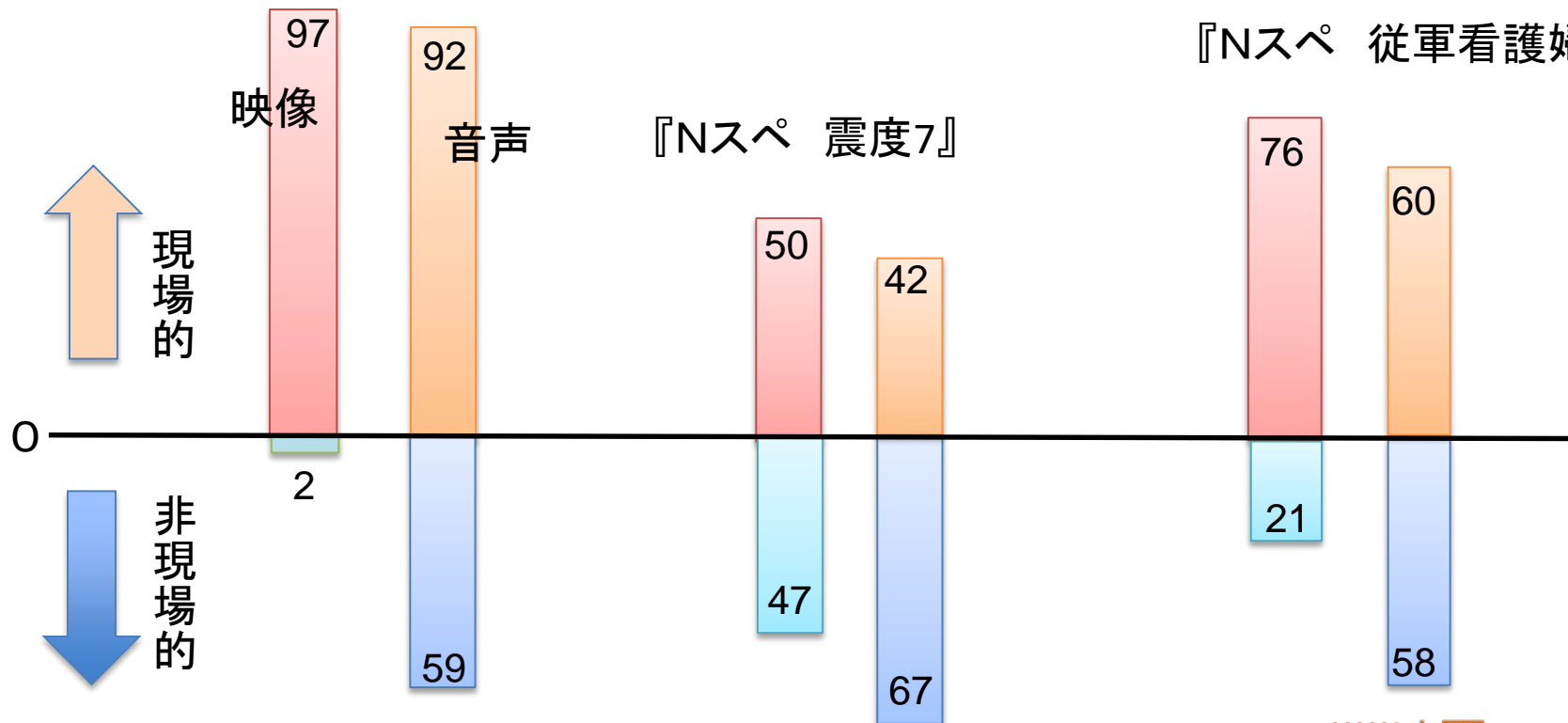
現場音 (AUB) と 非現場音 (CUD)



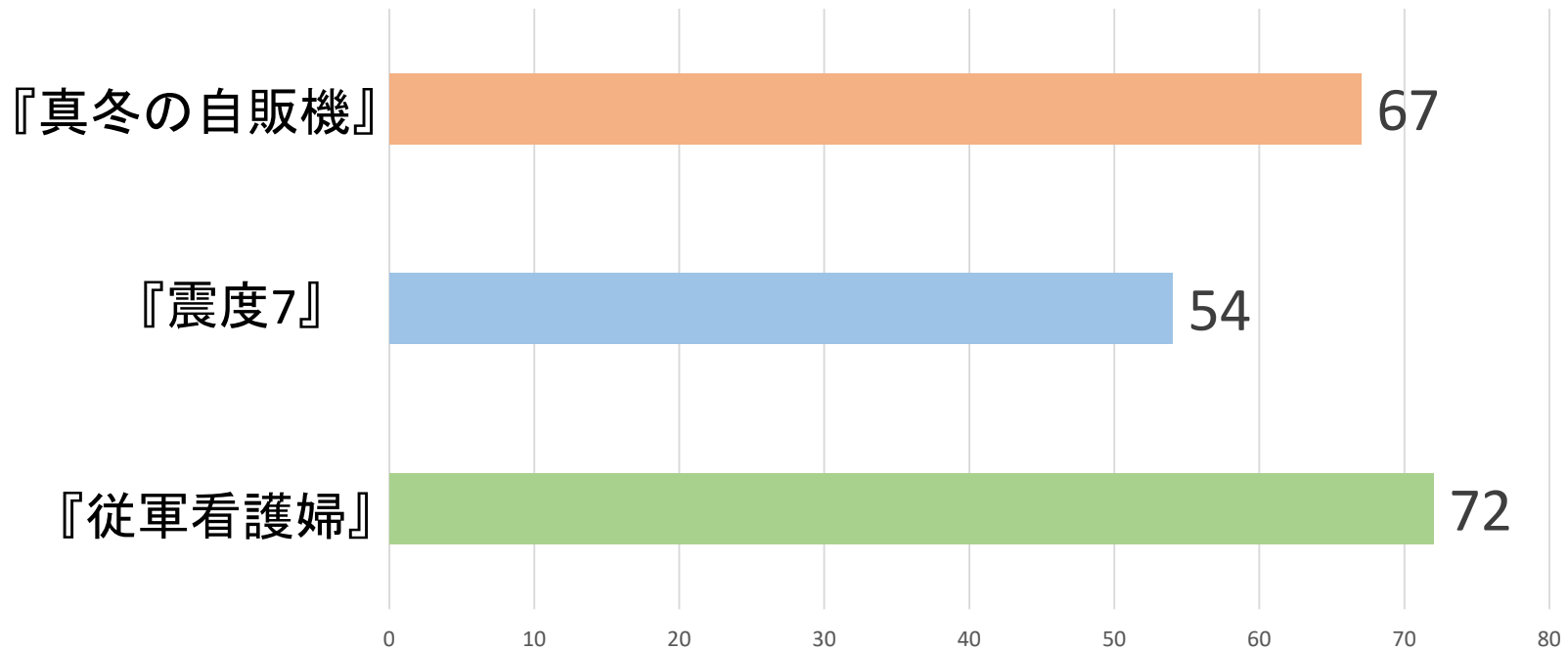
映像と音声の現場／非現場 (数字は%)

『72時間 真冬の自販機』

『Nस्प 従軍看護婦』



BUD、言語音声の被覆率(%)



ちなみに『ドキュメント72時間』は音声データも均一的である

(数字は上が尺数(秒)、下がテキスト実尺に対する比率)

テキスト	テキスト 実尺	現場の 物音 (A)	現場の 声 (B)	音楽等 (C)	ナレー ション (D)	現場音 尺 (AUB)	非現場 音尺 (CUD)	言語音 尺 (BUD)	有音部 総尺
『72時間』5本 平均	1470	1160	751	624	345	1291	792	1008	2880
	100	79	51	42	23	88	54	69	196
『72時間 真冬 の自販機』	1470	1350	737	685	362	1356	864	987	3134
	100	92	50	47	25	92	59	67	213

パート3:応用問題

『日本の素顔 日本人と次郎長』(1958)はなぜヒットしたか？

映像と音声のデータをどう生かすか？

二つの方向

1、マクロな研究:このデータを指標にして大きなテキスト群の特徴を別のテキスト群と比較しながらとらえる。

○歴史的な変容:たとえば70年代のTDと現代のTD

○作り手ごとの違い:たとえばNHKのTDと民放のTD、吉田直哉と牛山純一

○国内外の違い:日本のTDと諸外国のTD

2、ミクロな研究

このデータをもとに、テキスト内容に関する他の知見、また、テキスト外の知見も用いて、個々のテキストを精緻に読み解く。

『日本の素顔 日本人と次郎長』(1958年1月5日放送)

・・・吉田直哉がディレクターとして制作

○吉田によると、『日本の素顔』は1957年11月にベンチャー番組として当初12本の放送予定で始まった。

○しかし、この第8集『次郎長』のヒットによって「生き延び」、以後1964年まで306本を放送する長大なTDシリーズとなった※。

日本のTDの出発点となった最初のヒット作

⇒このテキストがなぜヒットしたかを考える。

視聴：『日本人と次郎長』

「なぜヒットしたか？」 作り手：吉田直哉の見解

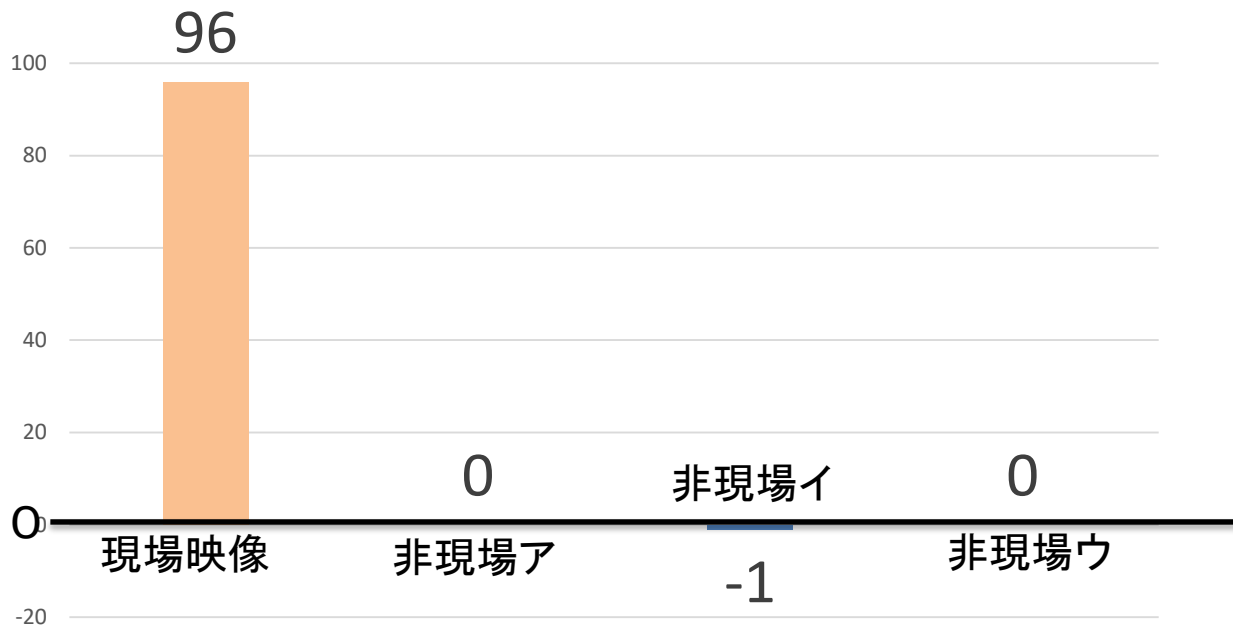
「現実のヤクザの一家の生態、つまり、極秘で開帳される賭博、親子の盃がための儀式、入れ墨をする情景、ヤクザ同士の喧嘩、その手打ち式などが**ムービーカメラで記録**され、公開されたのは初めてであったことから、この作品は、その**スクープ性**と、画面に登場する**実在のヤクザたち**を徹底的に批判した**(ナレーションの)****大胆さ**の故に記憶される運命となった」(吉田1977)。

吉田の見解を要約するとヒットの理由は
映像のスcoop性と、
ヤクザを真っ向から批判したナレーションの大胆さ

これが妥当かどうか、テキストの映像、音声についてのデータと、
当時の視聴者についての知見を材料にして考える。

まずデータから…

『次郎長』の映像(数字は%)



圧倒的に現場的

実尺1800秒中、非現場映像は
文書の接写2か所、計20秒だけ

現場映像を相対化する非現場映像がほとんどない(映像的に単層)

『次郎長』の音声

数字はテキスト実尺に対する比率(%)

	非言語音声	言語音声
↑ 現場的	0 (A音:現場の物音)	4 (B音:現場の声)
↓ 非現場的	99 (C音:音楽など)	36 (D:ナレーション)

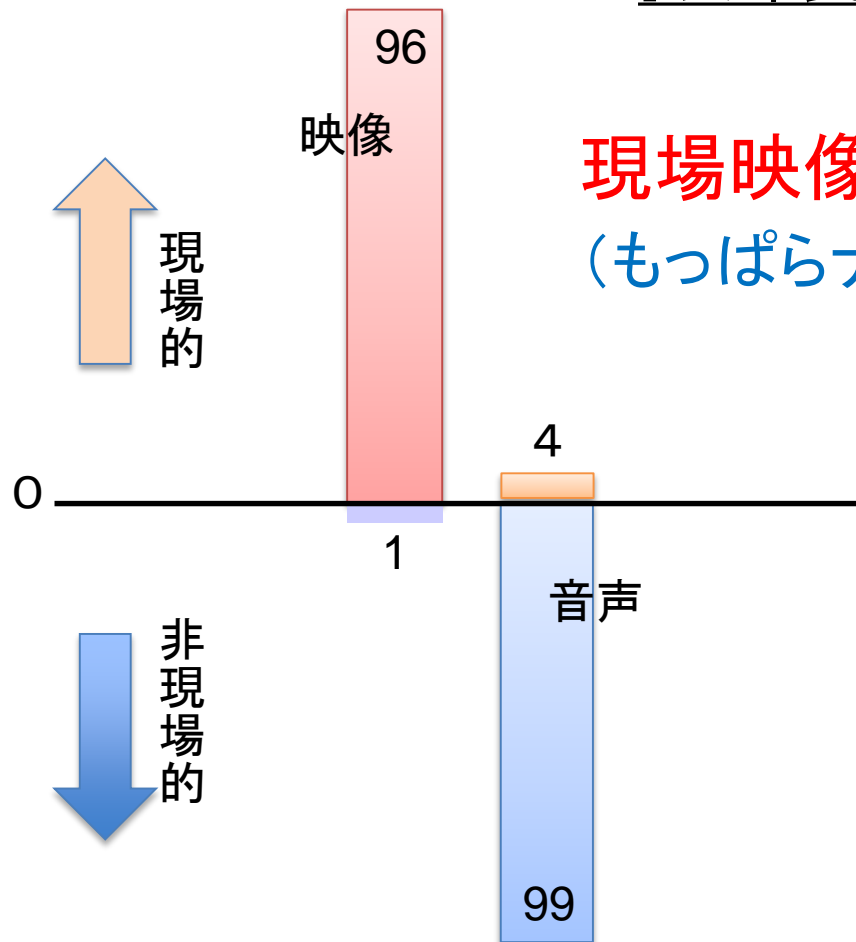
現場音声がほとんどない。

非言語音声、言語音声とも
圧倒的に非現場的

非現場音声のうち音楽(C音)は現場音声がなかったために大量に発生する無音部分を埋めるためのもので積極的な意味はない。

⇒音声ではナレーション(D音)が抜群の存在感※

『次郎長』の映像と音声の現場／非現場



現場映像を非現場音声
(もっぱらナレーション)で意味づけていく作り

当時の技術的制約

1950～60年代のTDでは、
ロケ現場での同時録音が困難であった。

カメラ自身が発するノイズが大きく、録音はカメラが回っていないときか、撮影中のカメラから離れた場所で行うしかなかった。

⇒音声はある程度非現場的であることを余儀なくされていた。

『次郎長』と並ぶ『日本の素顔』の古典的テキストである『奇病のかげに』は『素顔』の中では現場録音に注力したTD。それでも非現場音が現場音を大きく上回る。

『日本の素顔 日本人と次郎長』
(1958年1月5日放送)

非言語音声	言語音声
0 (A音:現場の物音)	4 (B音:現場の声)
99 (C音:音楽など)	36 (D:ナレーション)

『日本の素顔 奇病のかげに』
(1959年11月29日放送)

非言語音声	言語音声
2 (A音:現場の物音)	22 (B音:現場の声)
86 (C音:音楽など)	42 (D:ナレーション)

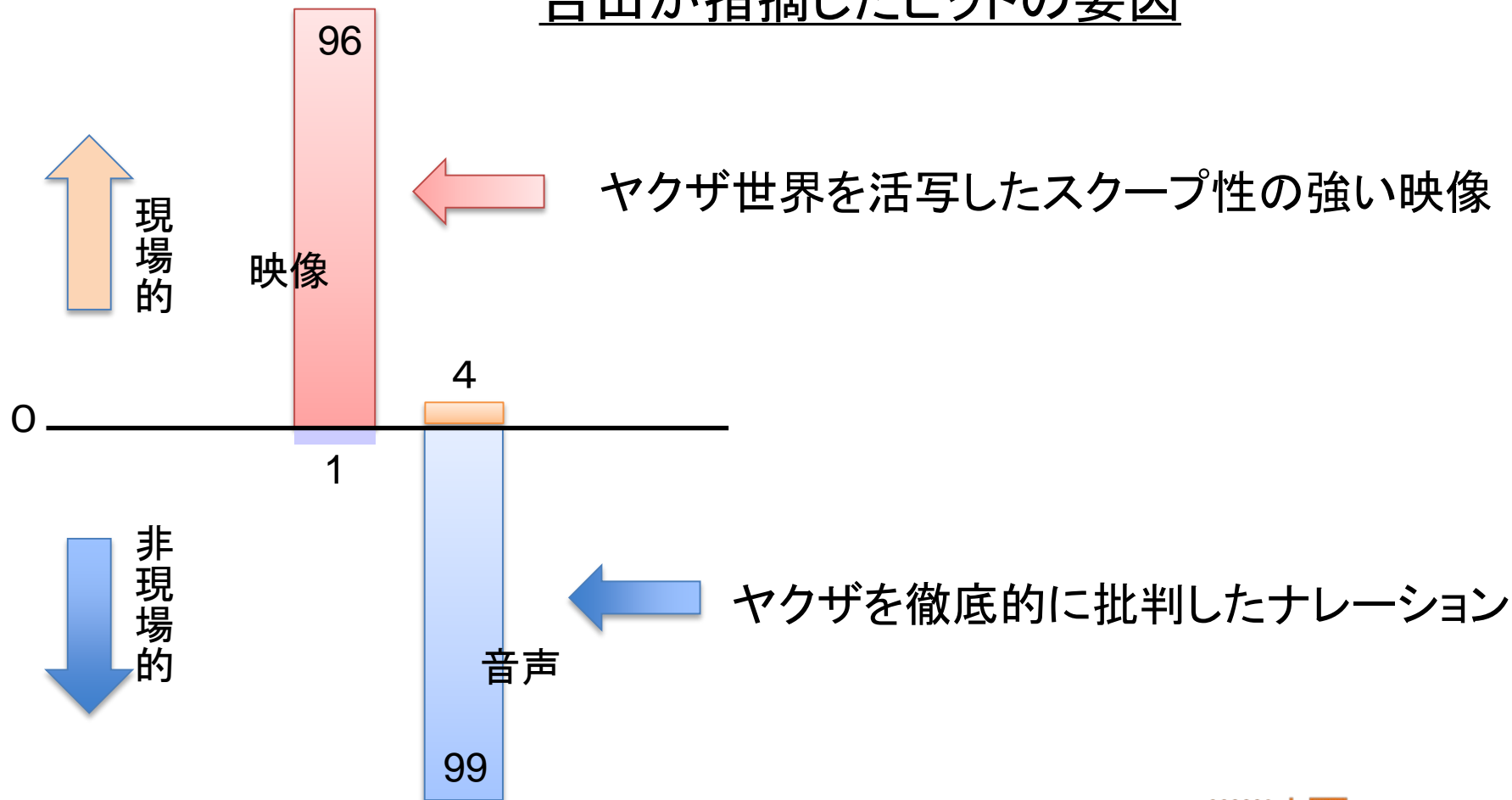
啓蒙的テキストが生まれやすい技術環境

丹羽美之氏の指摘:『日本の素顔』は作り手が近代主義の立場から、「社会を憂慮する」**啓蒙的ドキュメンタリー**である(丹羽2001、2003)。

テキスト内で「社会を憂慮する」のはもっぱらナレーション

同時録音が普及する以前のTDは、作り手が**ナレーションで、現場の生々しい現実を裁断する啓蒙的テキスト**が生まれやすい**技術環境**のもとにあった※。

吉田が指摘したヒットの要因

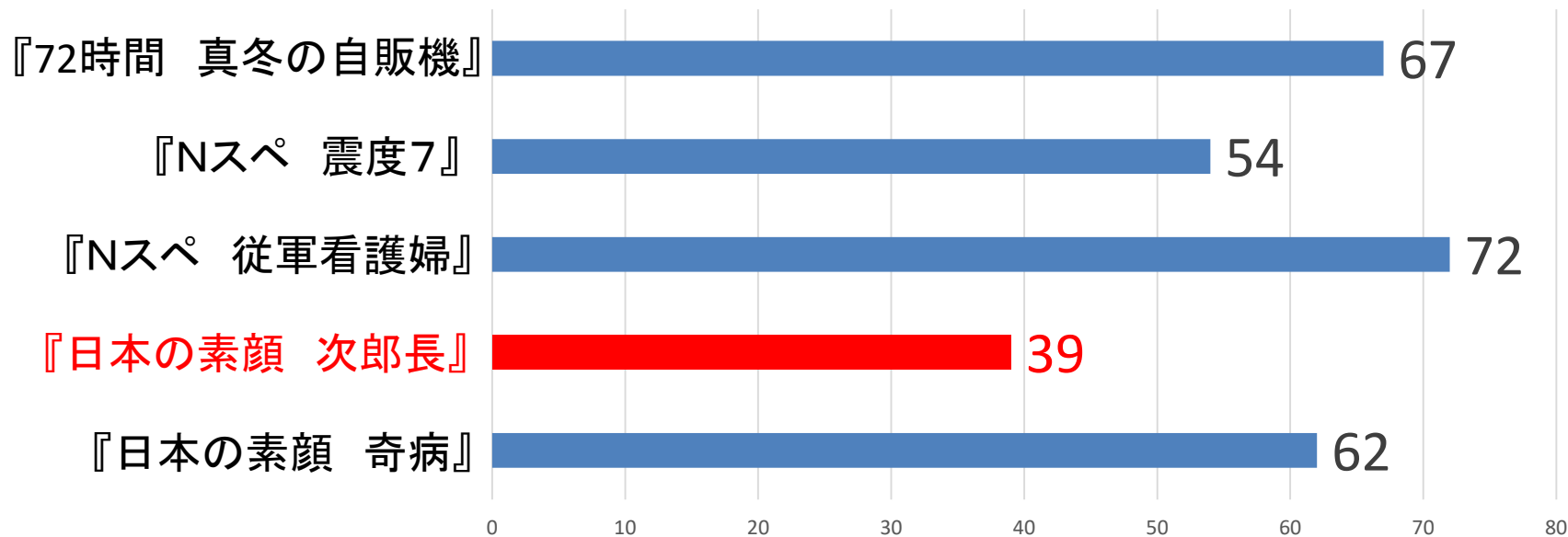


映像がヤクザ世界を活写したスクープ性豊かなものであることは間違いない。

「現実のヤクザの一家の生態、つまり、極秘で開帳される賭博、親子の盃がための儀式、入れ墨をする情景、ヤクザ同士の喧嘩、その手打ち式・・・」

データからわかるのは
これらの映像の多くが、
言語音声に随伴されず、いわば「映像単体」で展開していること。

言語音被覆率(BUD、数字は%)



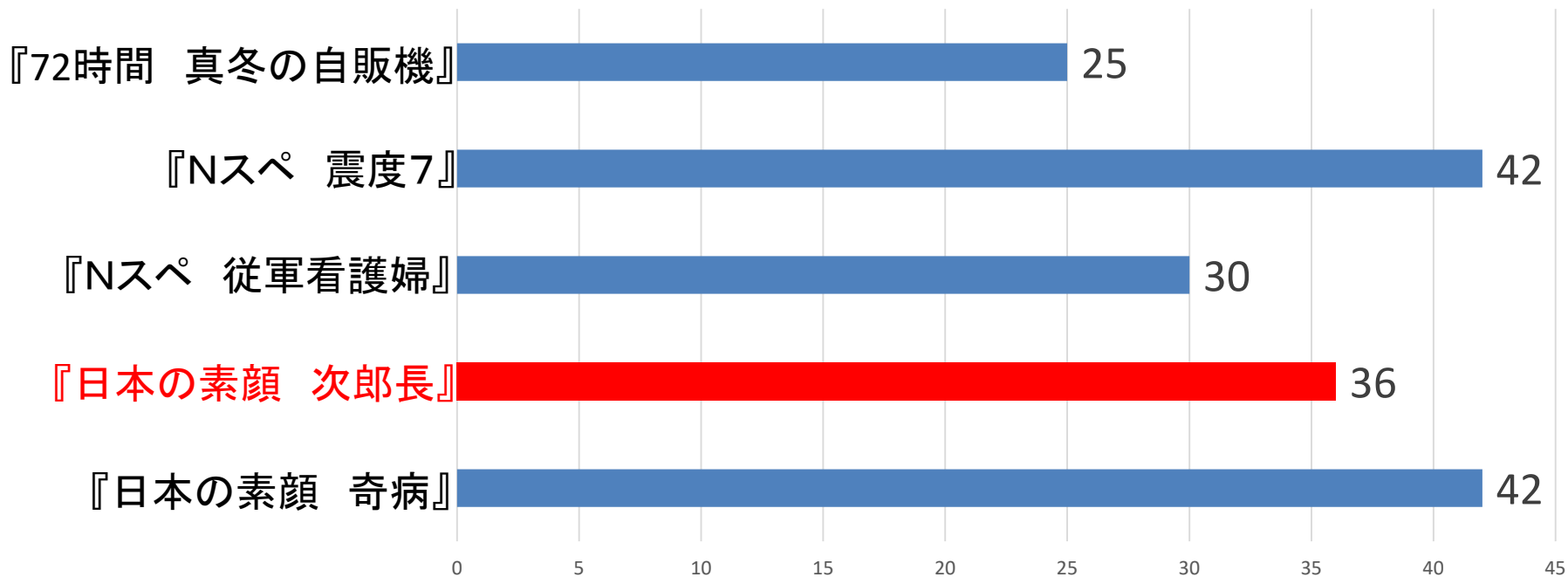
『次郎長』の映像は言語に随伴されている割合が低い
⇒61%は「映像単体」で展開している

『次郎長』がヒットしたということは、テキスト実尺の61%を占める「映像単体」に十分な魅力があったということではないか。

次に音声について考える。

『次郎長』の音声で抜群の存在感を持つナレーション
⇒これはどれだけヒットの要因となったのか。

ナレーション(D音)の量(数字は%)



『次郎長』のナレーションの量に目立った特徴はない※

『次郎長』に限らず「啓蒙的」と評される『日本の素顔』のナレーションはどれだけ視聴者にアピールしていたのか

データ以外の資料を探す

知識人・インテリ・・・「一億総白痴化」に与しない知的なテレビ番組として『素顔』を評価 ⇒ ナレーションがアピール

一方「大衆」にはナレーションは届いていなかった可能性※。

瀬川昌昭の証言

吉田と同じ『素顔』のディレクターであった瀬川昌昭は、1959年7月5日、自らの『素顔』のロケ中に、山谷の簡易旅館の「テレビ室」で労働者たちが、その時放送されていた『日本の素顔 右翼』を視聴しているのに出くわした※。
(瀬川1959『キネマ旬報』(239))

一緒に視聴しているうちに「どうも筋の運びと、観客の反応が
しっくり行かない」ことに気づく。

瀬川の証言

「例えば暴力右翼が、新橋ステージの平和大会に発煙筒を投げたりするとドッと来るが、肝心の右翼の現状分析のくだりに入って、解説に熱を帯びて来る頃には、みんな退屈そうな素振りを示し始める。・・(中略)・・要するにナレーション部分の理屈が難しすぎて、何のことかわからないのだ」

難しいナレーションは「大衆」には通じない。彼らはテキストを「音なし」で見ているとしてナレーションに頼りがちな『素顔』の制作を自戒(瀬川1959:128)。

瀬川の証言が示唆していること

『次郎長』のナレーションは「大衆」には届いていなかった
可能性

⇒ヒットの要因としては限定的だったのではないか

宮田の考察

『次郎長』のヒットの主因はヤクザ世界を活写した映像の魅力である。

ただしナレーション以外の音声が独特の役割を果たしている※

「ない音声」に注目する

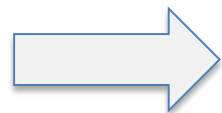
	非言語音声	言語音声
現場的	0 (A音:現場の物音)	4 (B音:現場の声)
非現場的	99 (C音:音楽など)	36 (D音:ナレーション)

吉田は、取材対象(ヤクザとその世界)が発する**現場音声**をほぼ**シャットアウト**している。

『次郎長』が描いているのは、現場映像と現場音声
がそろった**「生の」**ヤクザの**世界**ではない。

「なぜヒットしたか」考察まとめ

取材対象が発する自前の音声（現場音声）を封じることで、ヤクザ世界が持つある種の「生々しさ」を減らしたことが、『次郎長』がヒットした大きな要因となったのではないか※



『次郎長』というTDがヒットした最大の要因は、現場音声をほとんどシャットアウトして展開した現場映像の魅力である。