



---

Zur Problematik des automatischen Downmix  
von Mehrkanal auf Zweikanal

Helmut Wittek  
*Schoeps GmbH, Karlsruhe*

und Vincenz Riffeser

*vormals IRT, München*



- Was ist ein Downmix?
  - ITU
  - Matrix Encoder
  - Spatial Audio Coding
- Grundlagen zu Pegel und Abbildung
- Theorie und Praxis zu Klangfarbe und Räumlichkeit
  - Klangbeispiele (Aufnahme mit Hauptmikrofonen)
- OCT – ein Fallbeispiel
- Zusammenfassung



- Motivation
  - Notwendigkeit des automatischen Downmix
  - Kosten von zweifacher Regie, zweifacher Produktion, zweifacher Übertragung
- „Downmix-Kompatibilität“
  - einer Mischung
  - eines Hauptmikrofons
  - eines Effekts
  - ...
- Herangehensweise:
  - Identifikation einzelner Probleme
  - Finden von Parametern, die die Qualität des Downmix beeinflussen
  - Isolieren und Darstellen dieser Parameter
  - Diskussion über Lösungen
  - Möglichkeiten der theoretische Analyse: Pegel, Abbildung

- "Handmix"
- Transaural-Downmix
- ITU-Downmix
  - $L_t = L + k_1 * C + k_2 * L_s$
  - $R_t = R + k_1 * C + k_2 * R_s$   
mit  $k_1 = 0,71$  und  $k_2 = 1/0$
- Matrix Encoder (Dolby Pro Logic II)
  - aktiver/passiver Encoder
  - dynamische Signalanalyse
  - aktive Pegelsteuerung
  - spezifische Eigenschaften
    - $\Phi(L/L_s) = 90^\circ$
    - geringer Crosstalk  $RS \rightarrow L_t$
- spezielle Encoder
  - "Spatial Audio Coding": Nur 2 Parameter, die zur Codierung
  - Downmix-Processor ? (Ext)

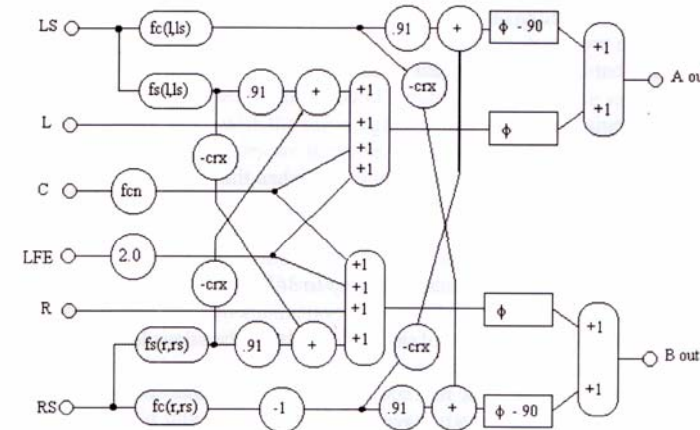
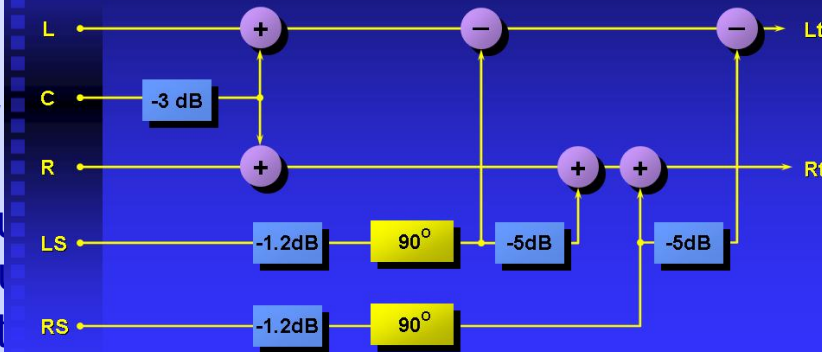


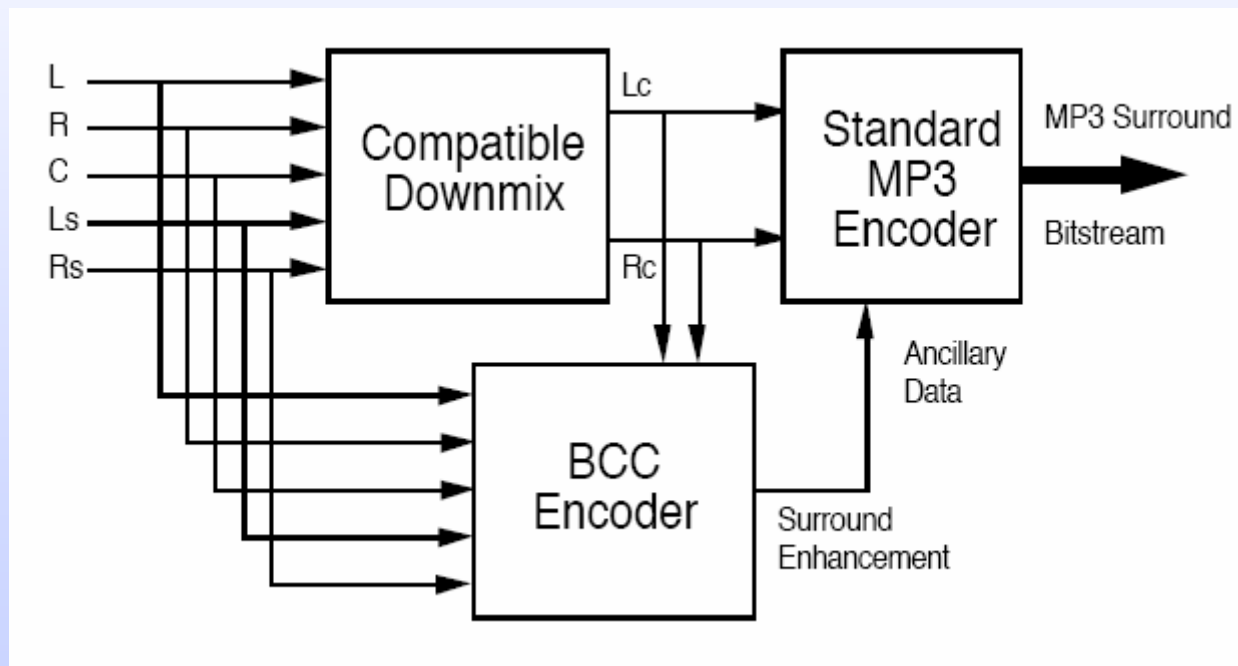
Figure 25. The new active encoder diagram

## Dolby Pro Logic II Encoder





- Spatial Audio Coding (Binaural Cue Coding)
  - z.B. mp3 Surround
  - Jeder beliebige Downmix möglich
  - Binaural Cues: ICTD, ICLD, ICC



aus: Herre et al. (FhG): MP3 Surround



## Hörbeispiel „Diana Krall“

No.

- Höreindruck

- *5.1 Surround*

1

- Stereo-Handmix

4

- ITU-Downmix

3

- Vergleich Stereoverision – ITU

- Handgemachte  
Zweikanalfassung

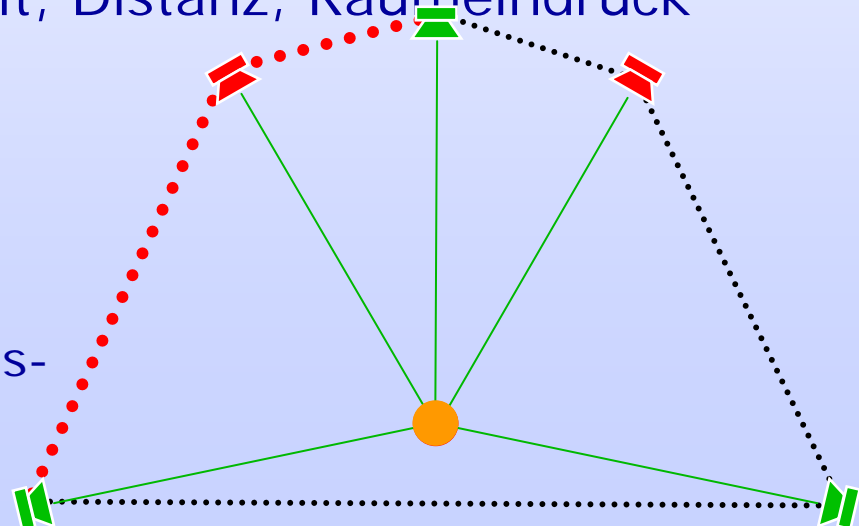
2

- ITU Downmix



## Grundsätzliche Probleme beim Downmix:

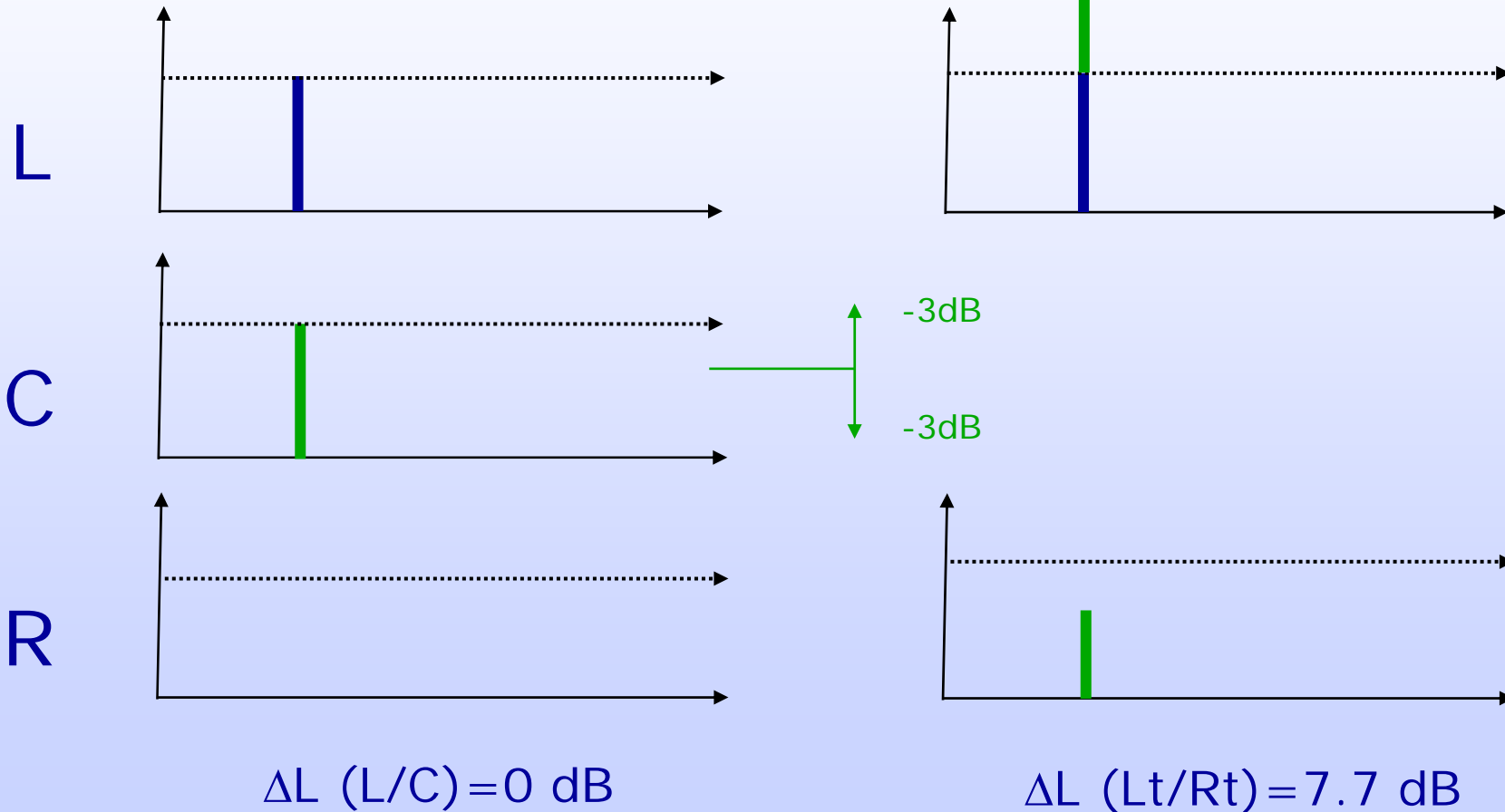
- Übersetzung des Konzepts (Dramaturgie)
- Zusammenklappen der Lautsprecherbasen erzeugt Dichte. Störend z.B. für Reflektionen
- Ohr bildet Summenpegel anders (3dB) als die elektronische Addition (6dB)
- Abbildung von Räumlichkeit, Distanz, Raumeindruck
- Pegelprobleme
  - vordere Basis (L-C)
  - seitliche Basis (L-LS)
- Abbildungsverzerrungen:
  - Pegelpanning/ Intensitätsstereofonie
  - Laufzeitstereofonie





## 5.1

## Downmix

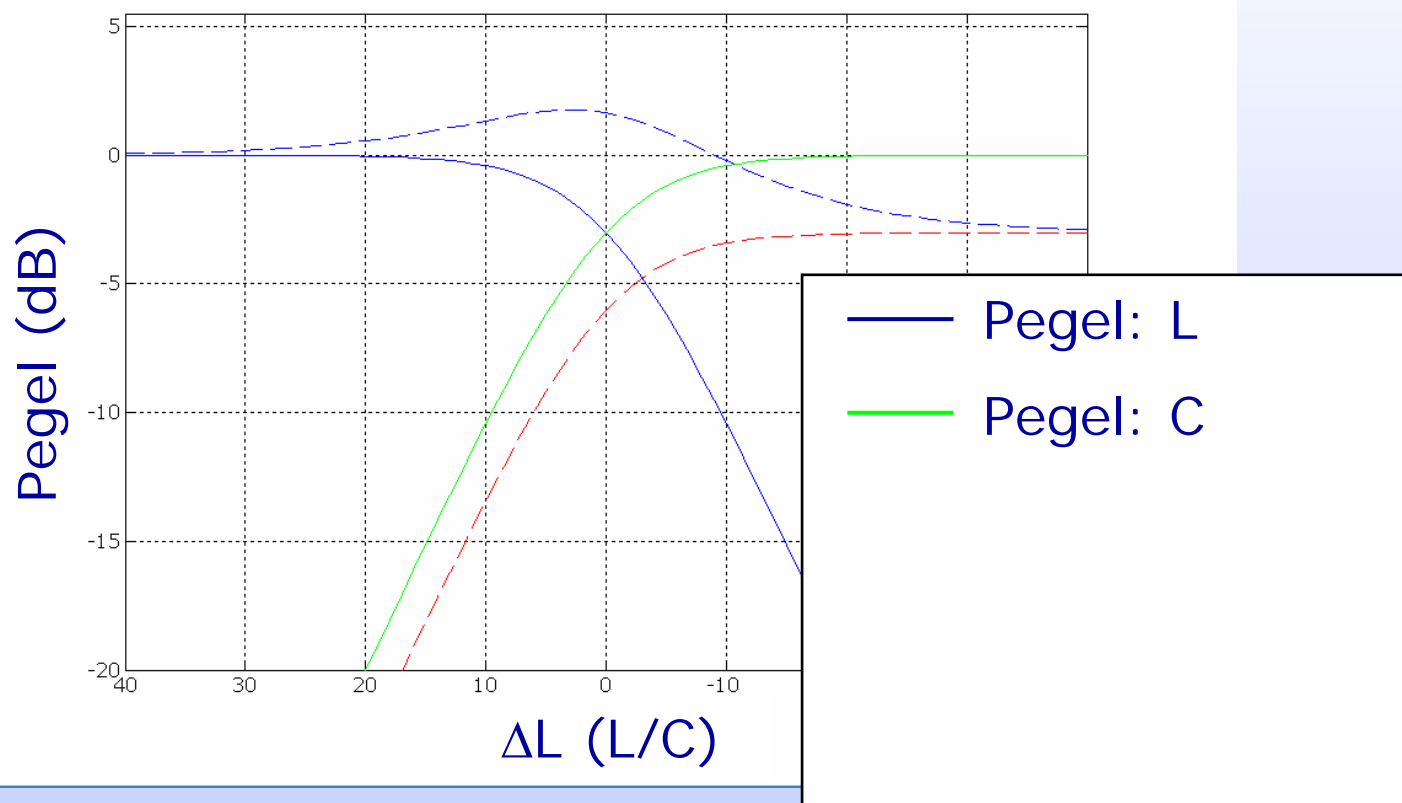




## Pegelpanning: Pegelkonstanz

- Theorie:
1. Pegel bei Pegelpanning über L-C : 2.3 dB - Fehler
  2. Pegel bei Pegelpanning über L-LS: 3dB - Fehler

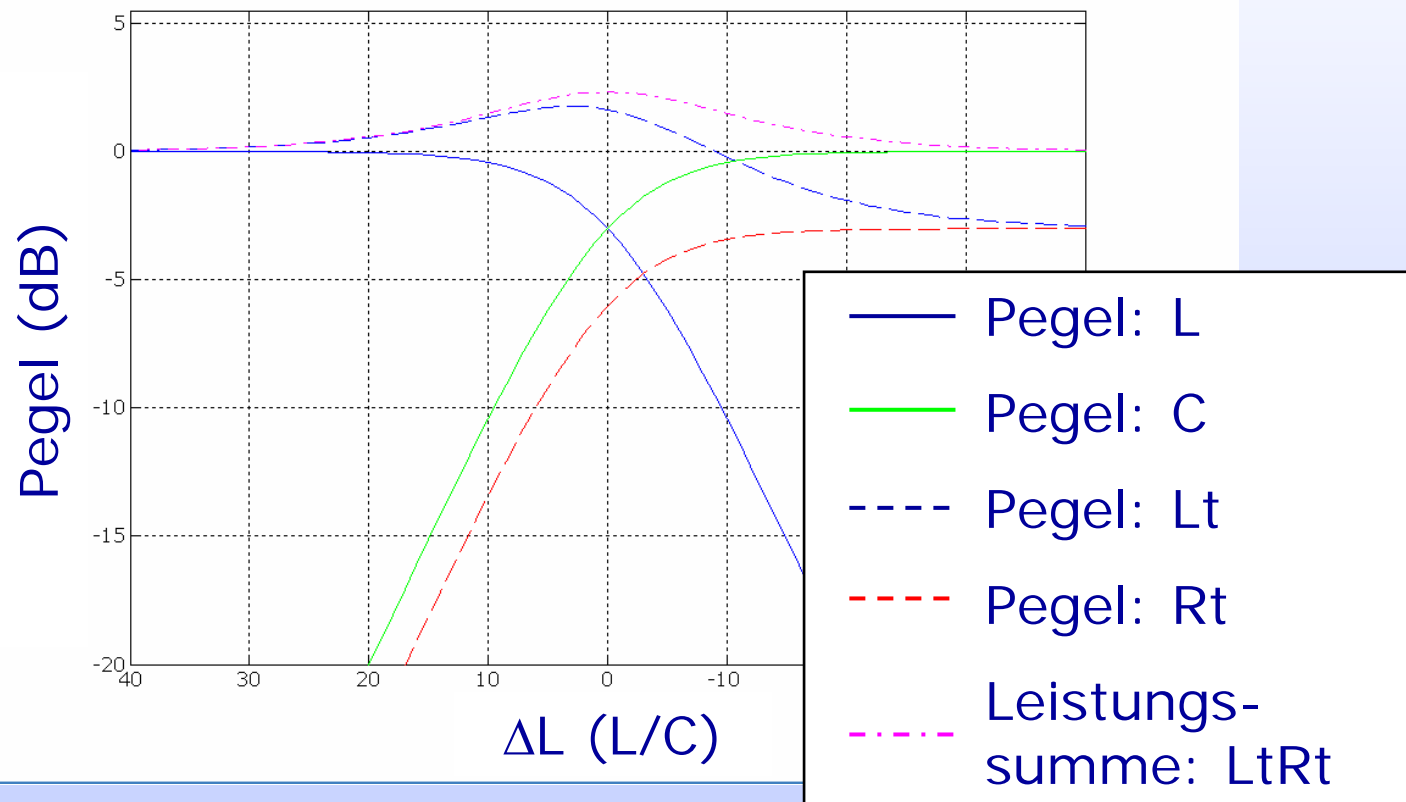
*Panning zwischen L und C*



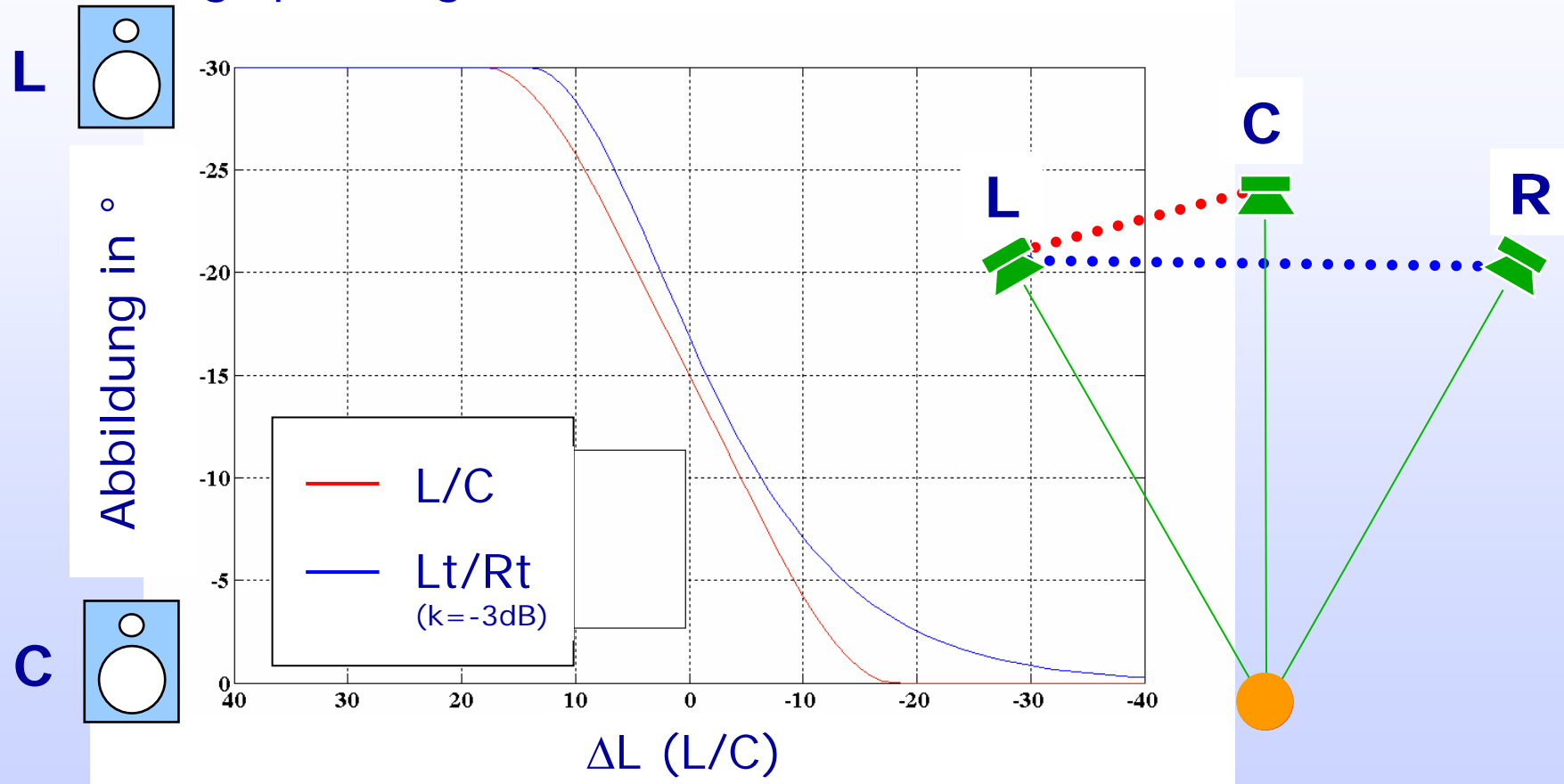
# Pegelpanning: Pegelkonstanz

- Theorie:
1. Pegel bei Pegelpanning über L-C : 2.3 dB - Fehler
  2. Pegel bei Pegelpanning über L-LS: 3dB - Fehler

## *Panning zwischen L und C*



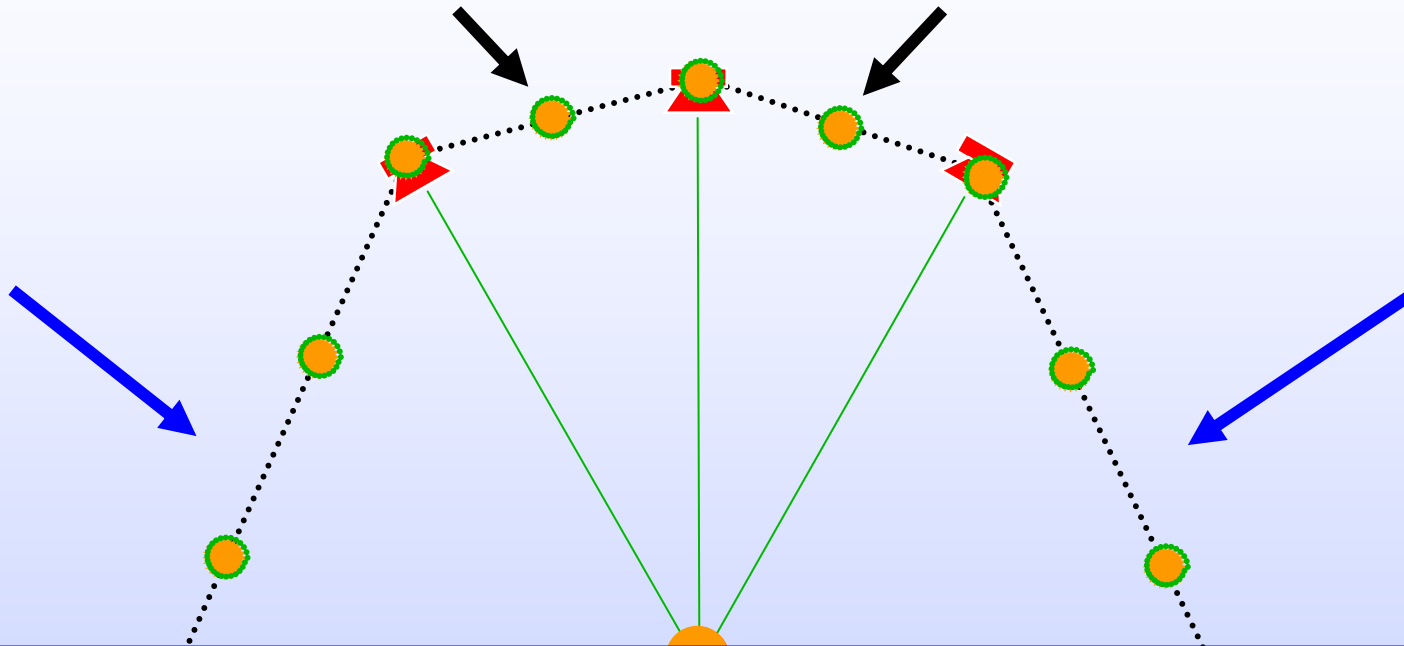
## Pegelpanning: vor und nach dem Downmix



Pegelpanning: Abbildung wird **breiter**

No. Pegelkonstanz bei Pegelpanning über L-C-R-LS-RS: Rauschen

5  
6



Abhilfe für Pegelproblem L/LS und R/RS:

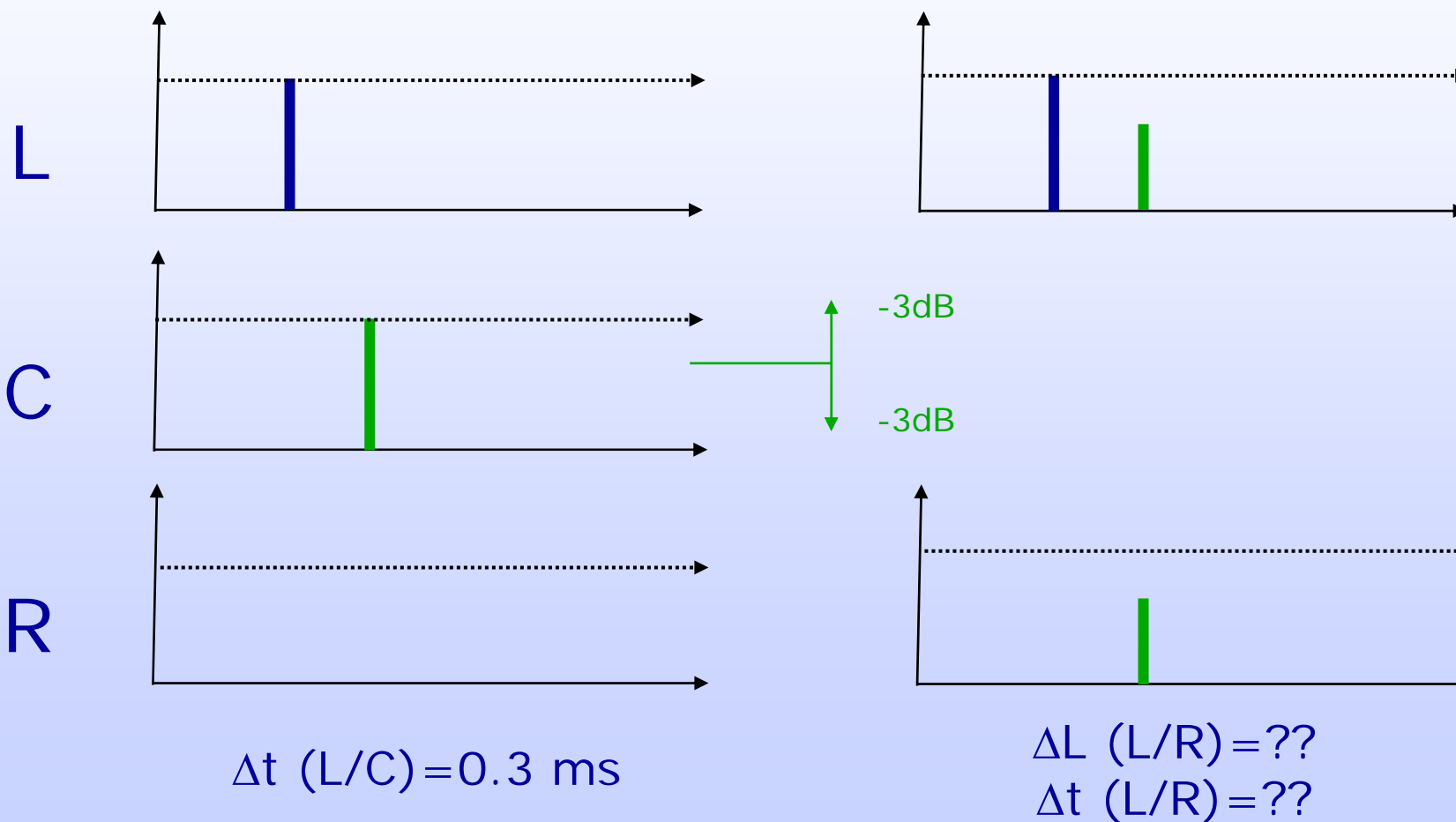
beim Downmix 90°-Phasendrehung (Logic7, Dolby)  
zwischen vorne und hinten (auf LS+RS)

→ Energie bei Downmix zwischen L/LS und R/RS konstant



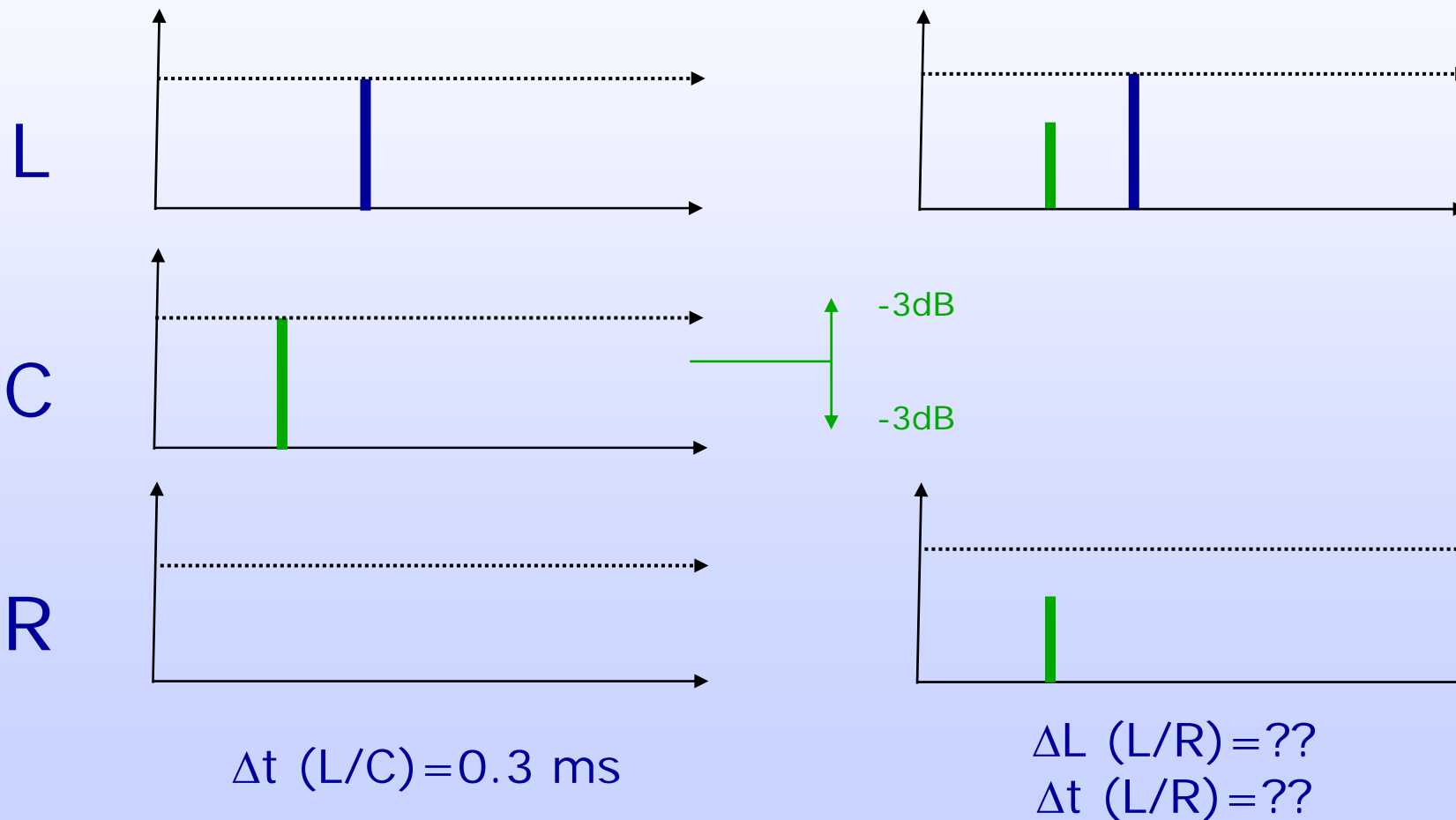
## 5.1

## Downmix



## 5.1

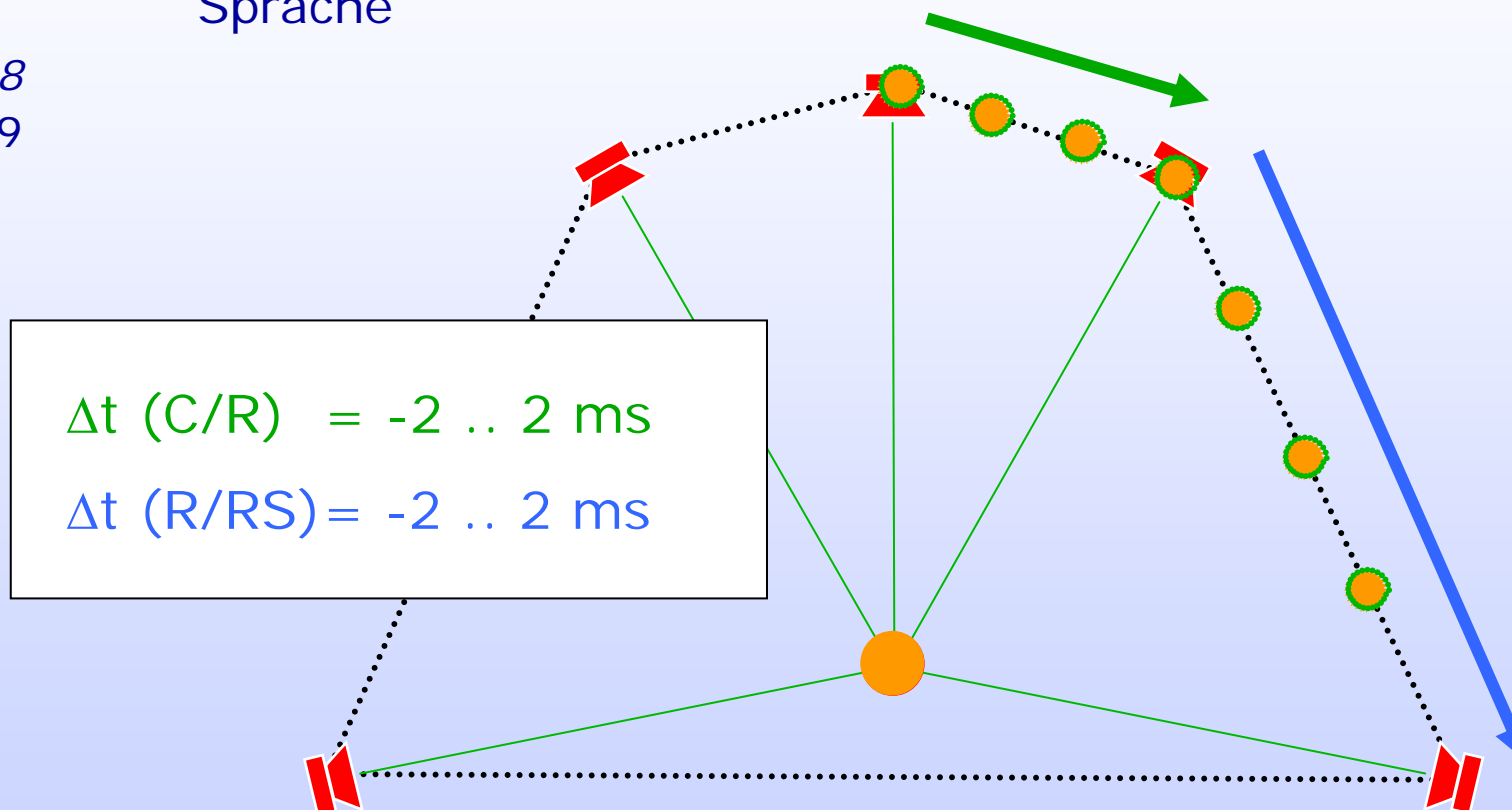
## Downmix



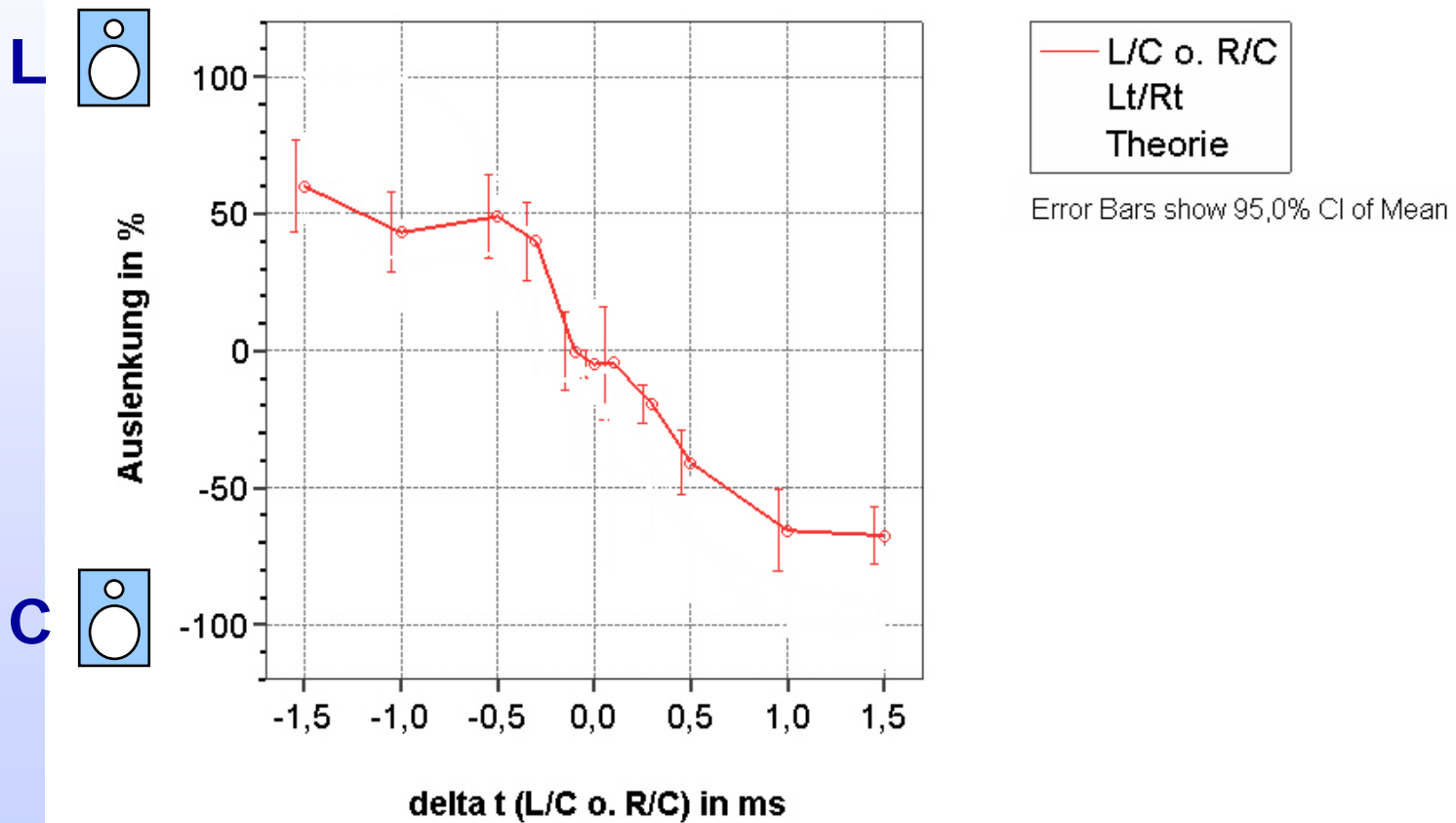
No.

Abbildung und Klangfarbe bei Laufzeit zwischen über L-C und C-R:  
Sprache

8  
9

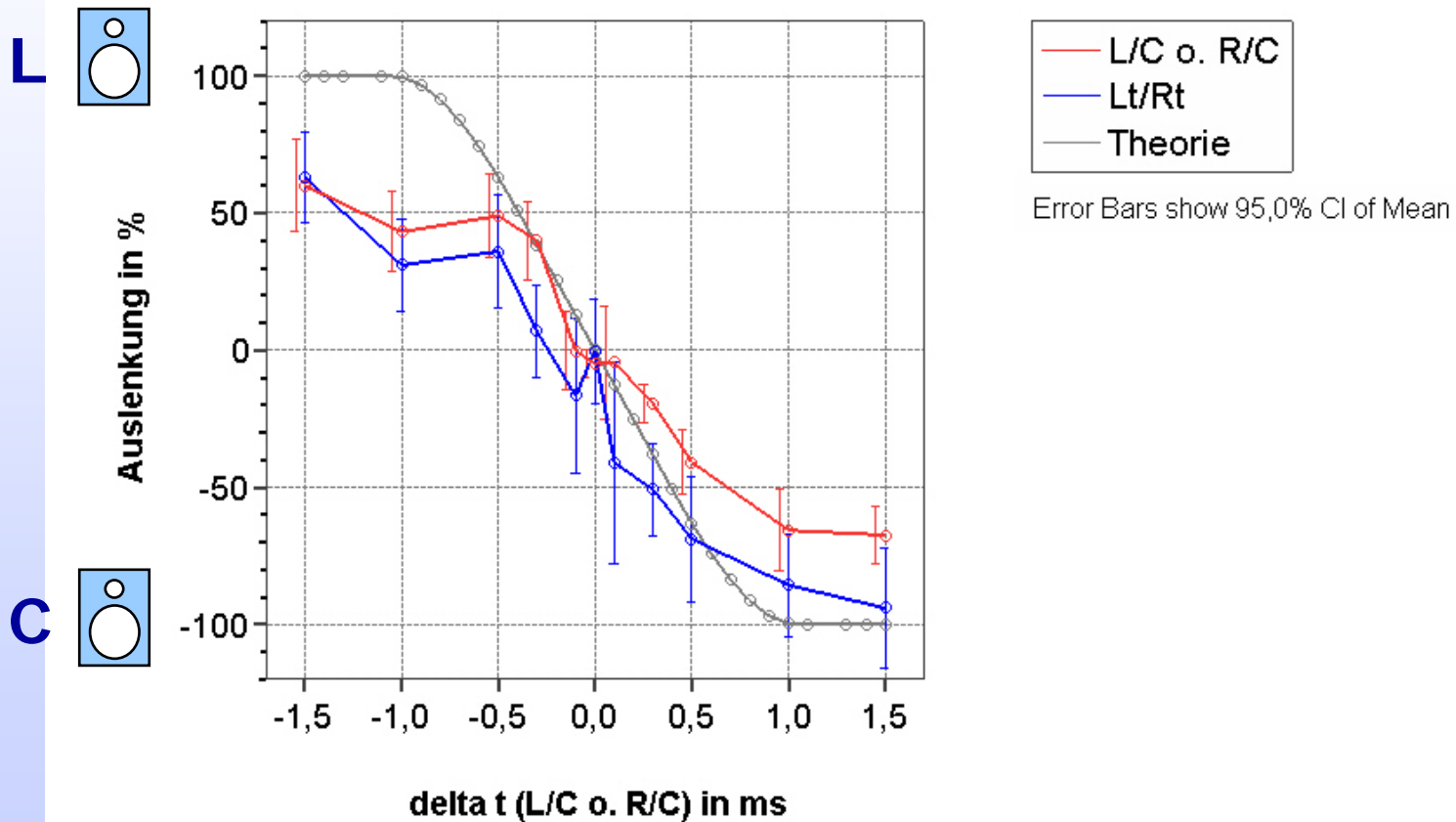


# Experiment: $\Delta t$ -Abbildung nach dem Downmix



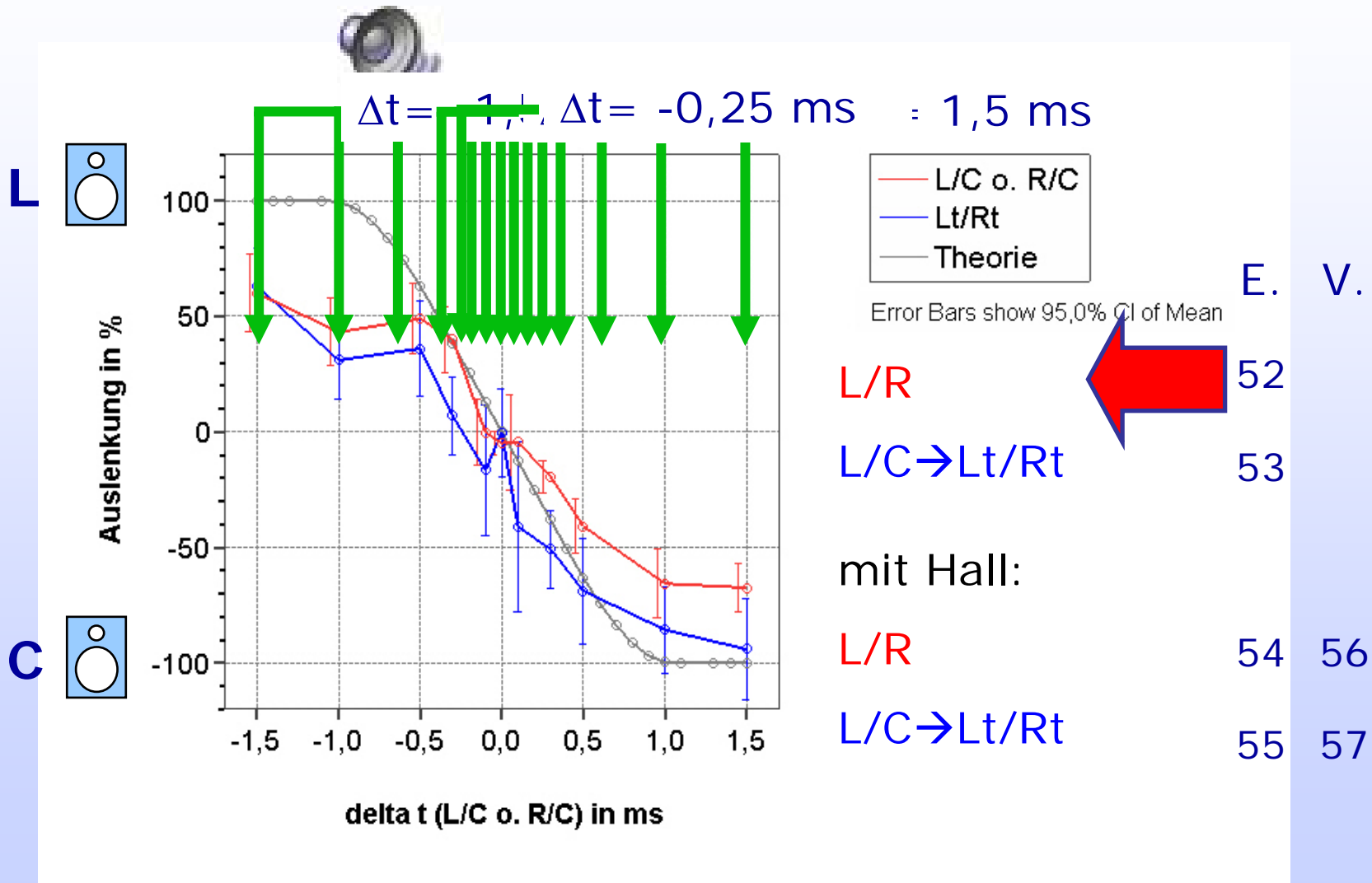


# Experiment: $\Delta t$ -Abbildung nach dem Downmix

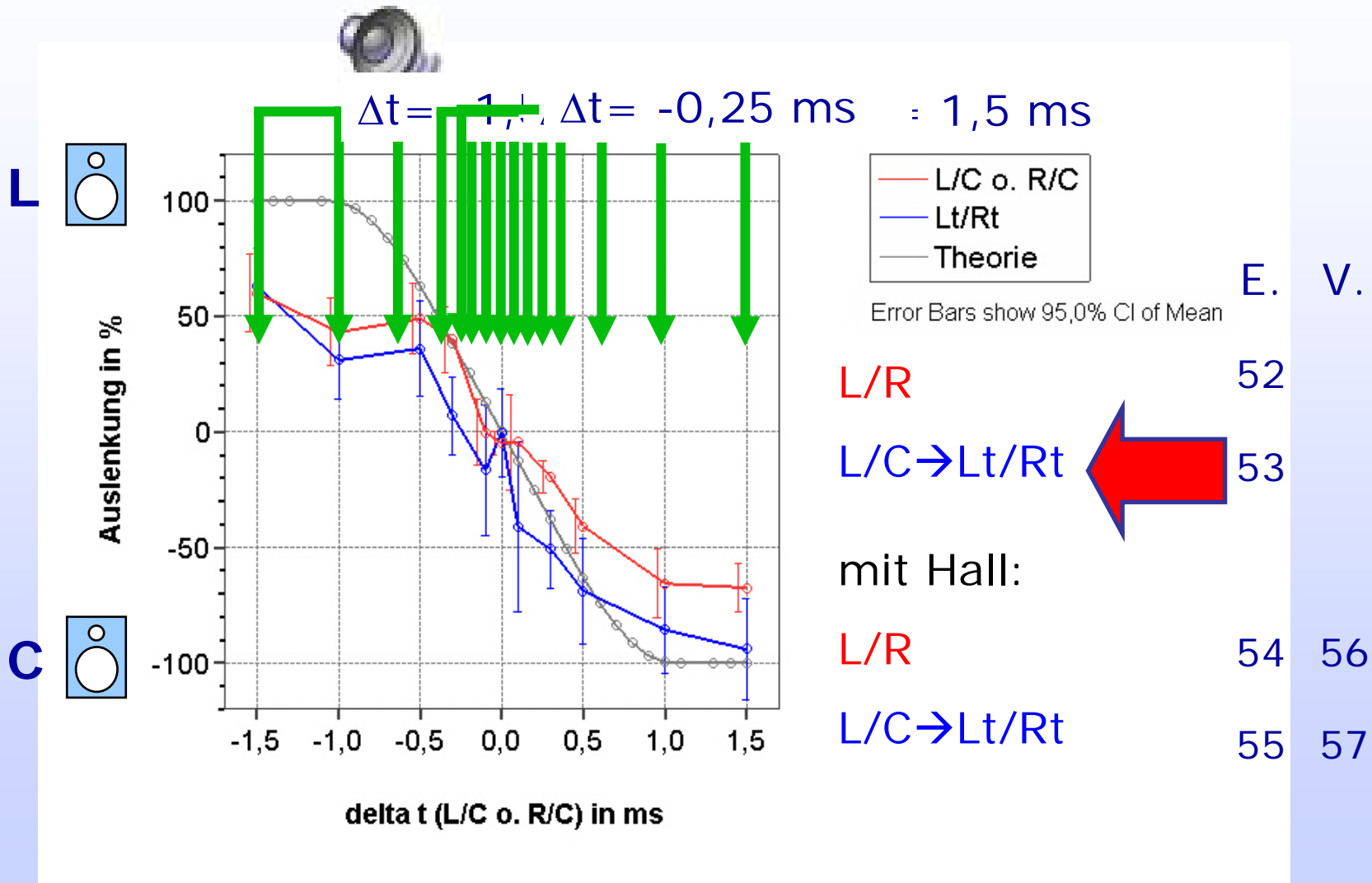


Laufzeitpanning: Abbildung wird **enger**

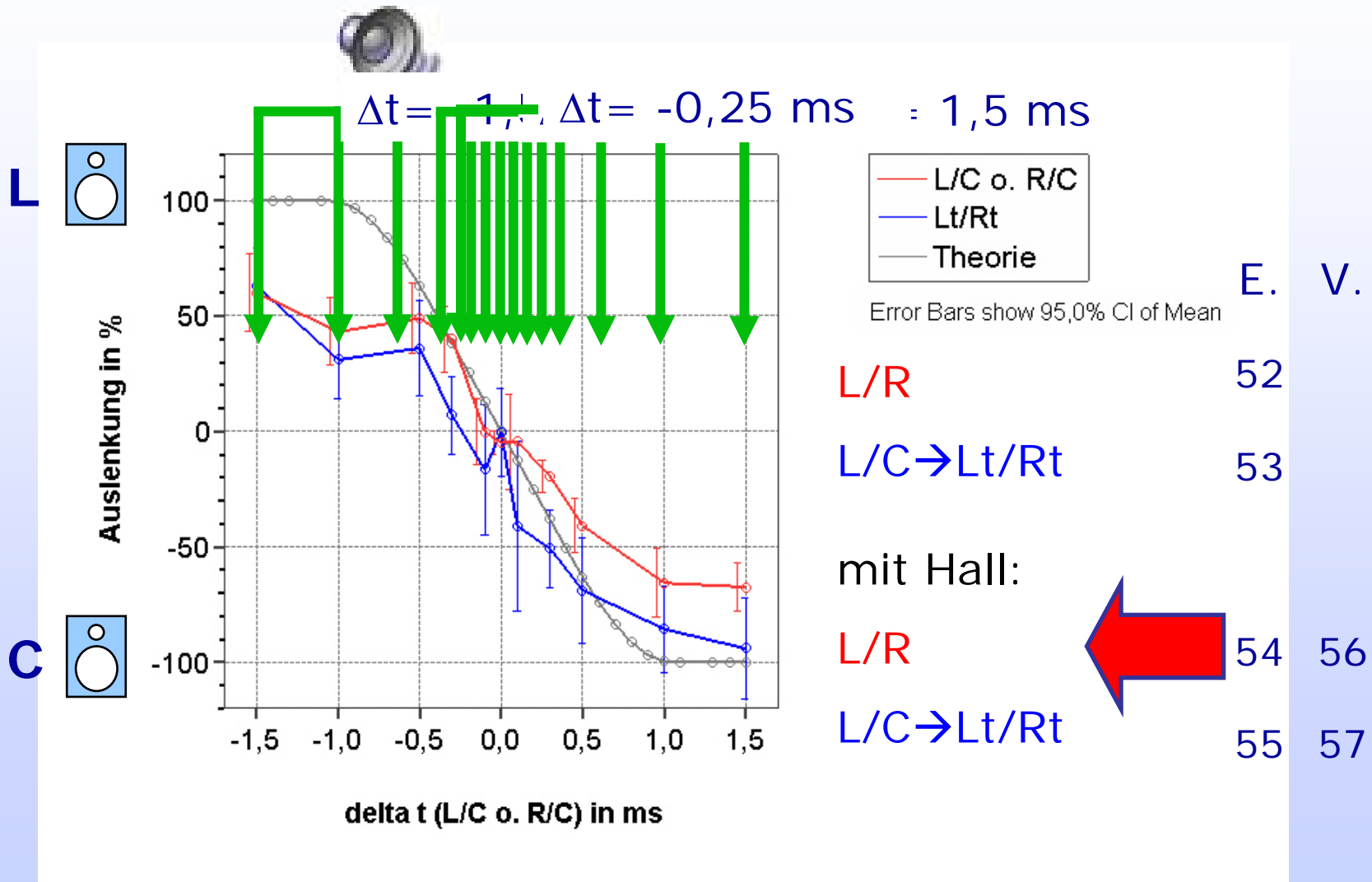
# Experiment: $\Delta t$ -Abbildung nach dem Downmix



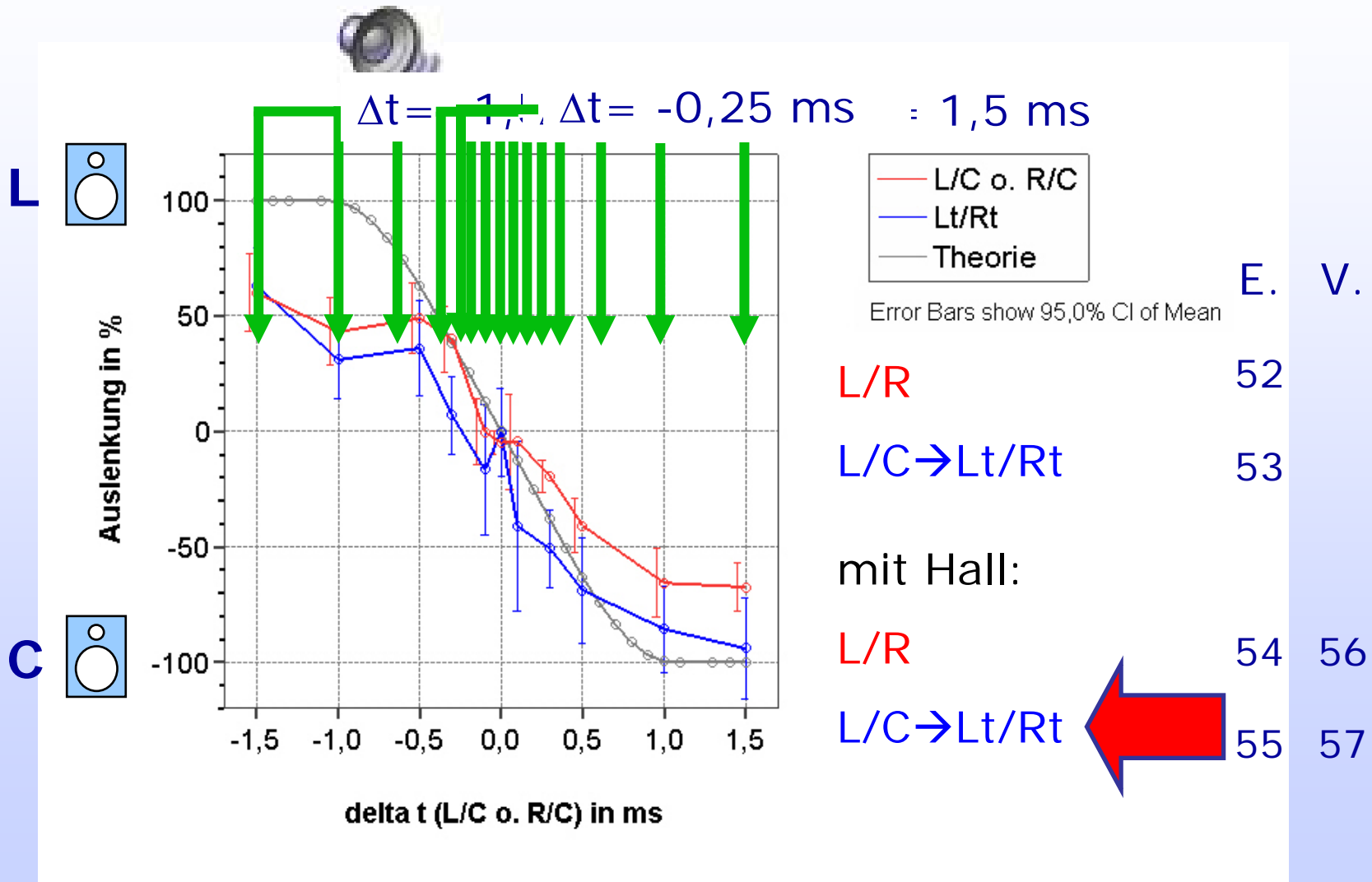
# Experiment: $\Delta t$ -Abbildung nach dem Downmix



# Experiment: $\Delta t$ -Abbildung nach dem Downmix

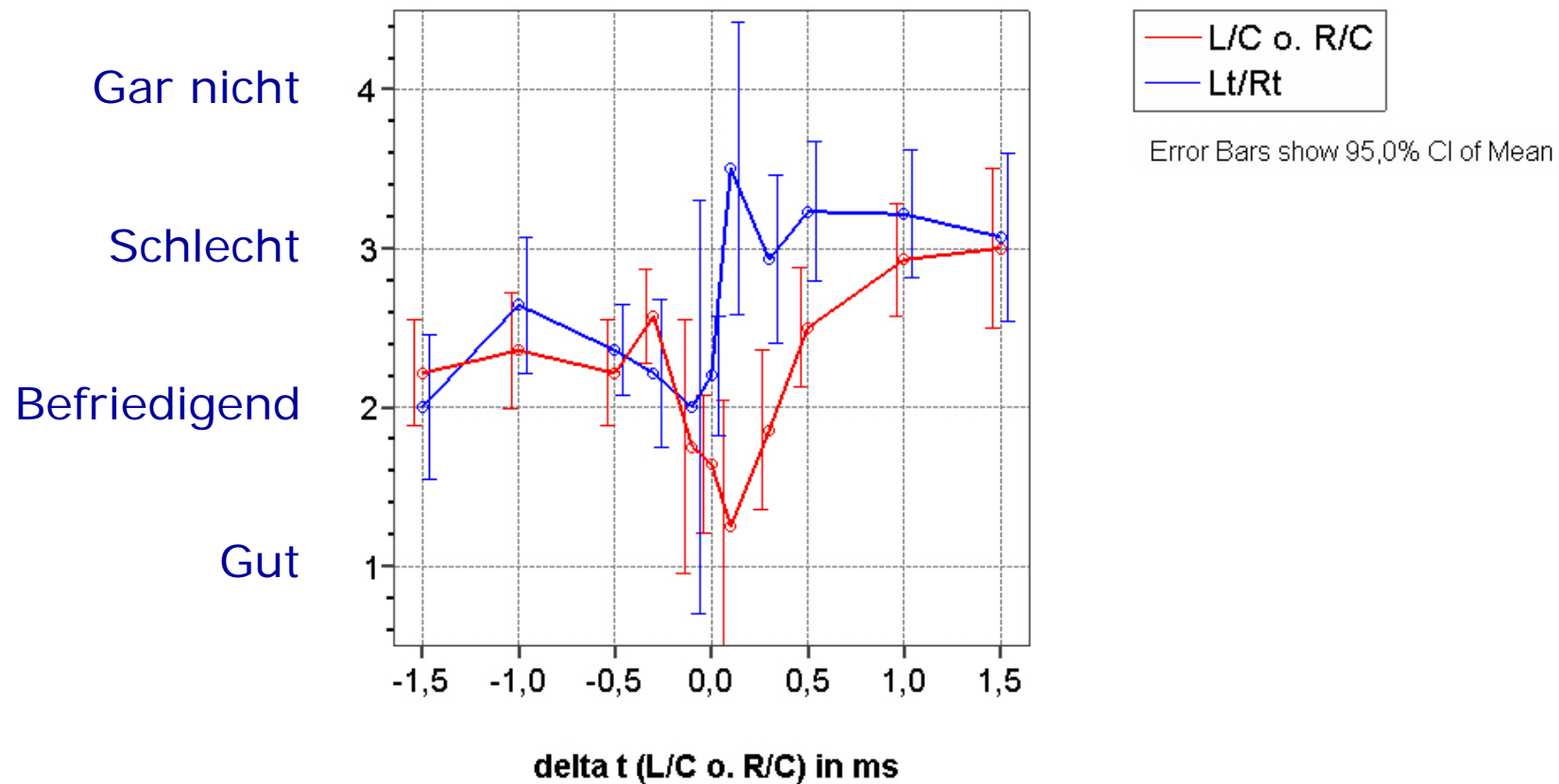


# Experiment: $\Delta t$ -Abbildung nach dem Downmix





## Wie gut ist die Quelle lokalisierbar?

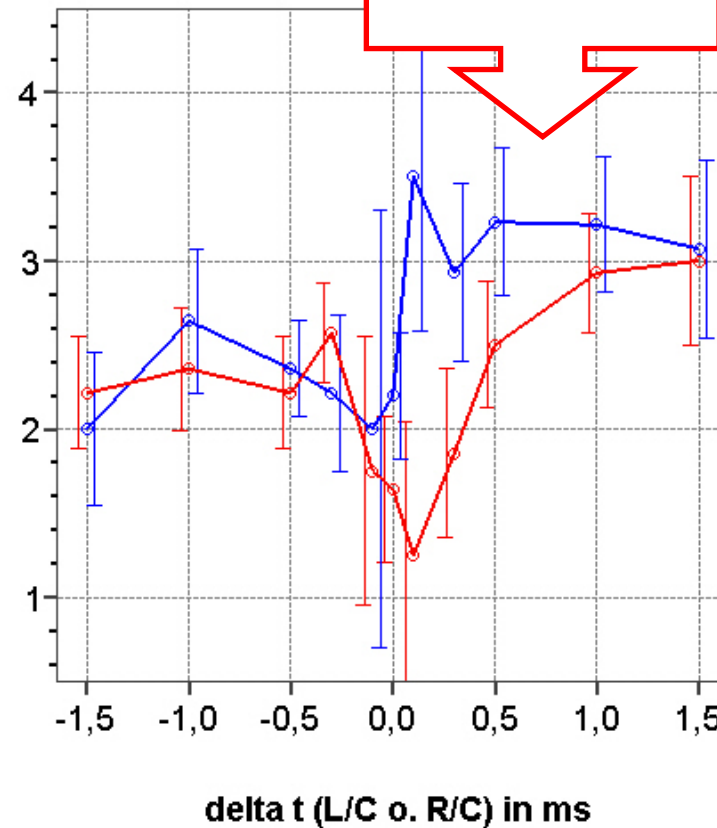




Wie gut ist die Quelle lokalisiert

Hören

Gar nicht  
Schlecht  
Befriedigend  
Gut



— L/C o. R/C  
— Lt/Rt

Error Bars show 95,0% CI of Mean

L/R

L/C → Lt/Rt

mit Hall:

L/R

L/C → Lt/Rt

E. V.

52

53

54

55

56

57

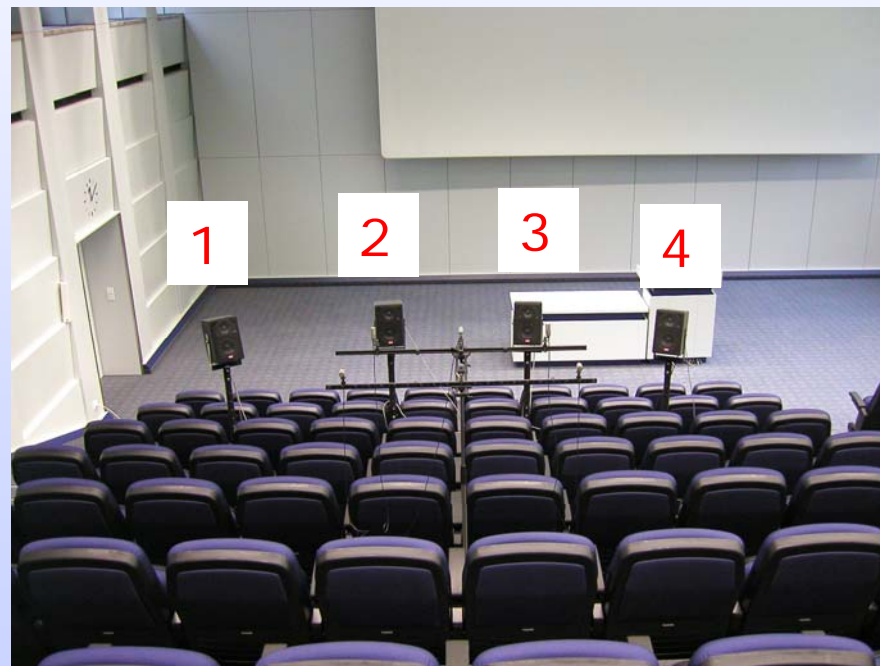


- Bilder Hörsaal und Skizzen der Anordnungen
  - Sprechstimme im raR
  - Klavier mit „The Grand“
- } Hauptsächlich trocken,  
reflexionsarm für die Faltung
- Erklärung des Versuchsaufbaus im Hörsaal zur Messung
  - Entstehung der Signale (Faltung...)
    - Lautsprecher, Impulsantwort, Messprogramm, Faltung
  - Vorstellung der verwendeten Mikrofonierungen
    - Vorspielen der verschiedenen Klänge (Sprache)





- Hörsaal im IRT
- 4 Schallquellen (Lautsprecher Nr.1 bis 4)
- Verschiedene Mikrofontechniken zur Aufnahme des „Quartettes“





- Mikrofontechniken exemplarisch für Parameter (Korrelation, Laufzeit, Pegel, Crosstalk, ...)
- Quellsignal für Schallquelle: weibliche Sprache, Klavier
- Signalgewinnung durch Impulsantwortmessung + Faltung (Qualitätsgewinn!)





Trockenes Signal für Schallquelle:

No.

- Sprachaufnahme im reflexionsarmen Raum  
Achten Sie *nicht* auf den Inhalt!



10

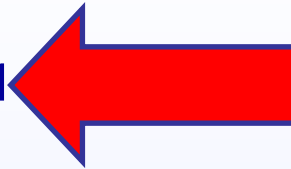
No.

## Kennenlernen Mikrofonklänge



12

- ORTF Pos. 1,2,3,4,3,2,1  
Positionswechsel

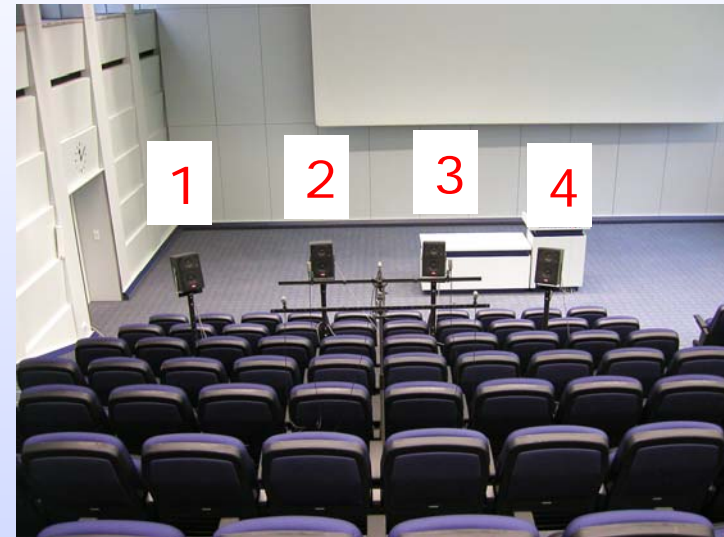


13

14

15

16

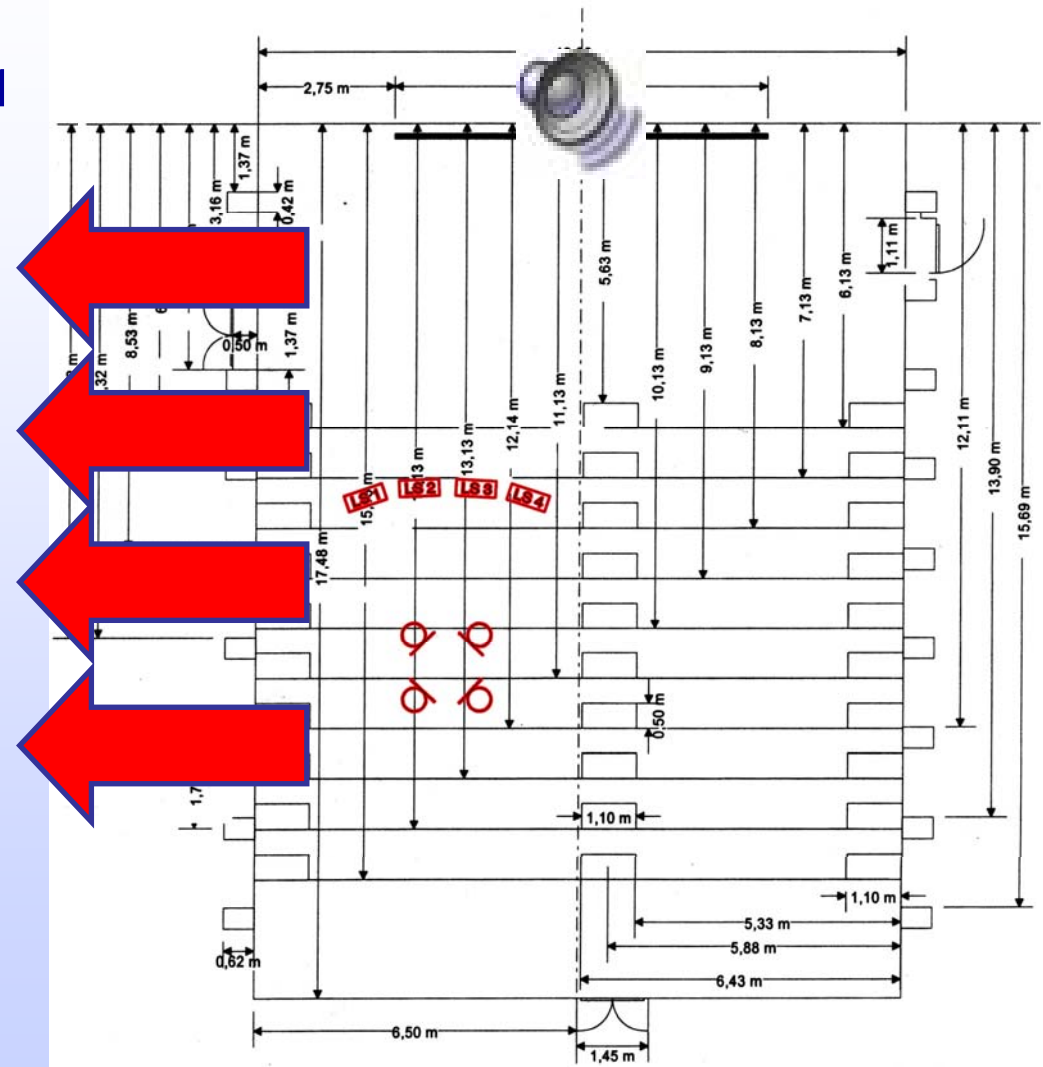


No.

## Kennenlernen Mikrofonklänge



- 12 • ORTF Pos. 1,2,3,4,3,2,1  
Positionswechsel
- 13 • IRT Kreuz klein  
Kantenlänge 30 cm
- 14 • IRT Kreuz groß  
Kantenlänge 99 cm
- 15 • OCT 1  
70/90/8
- 16 • OCT 2  
90/120/40



No.

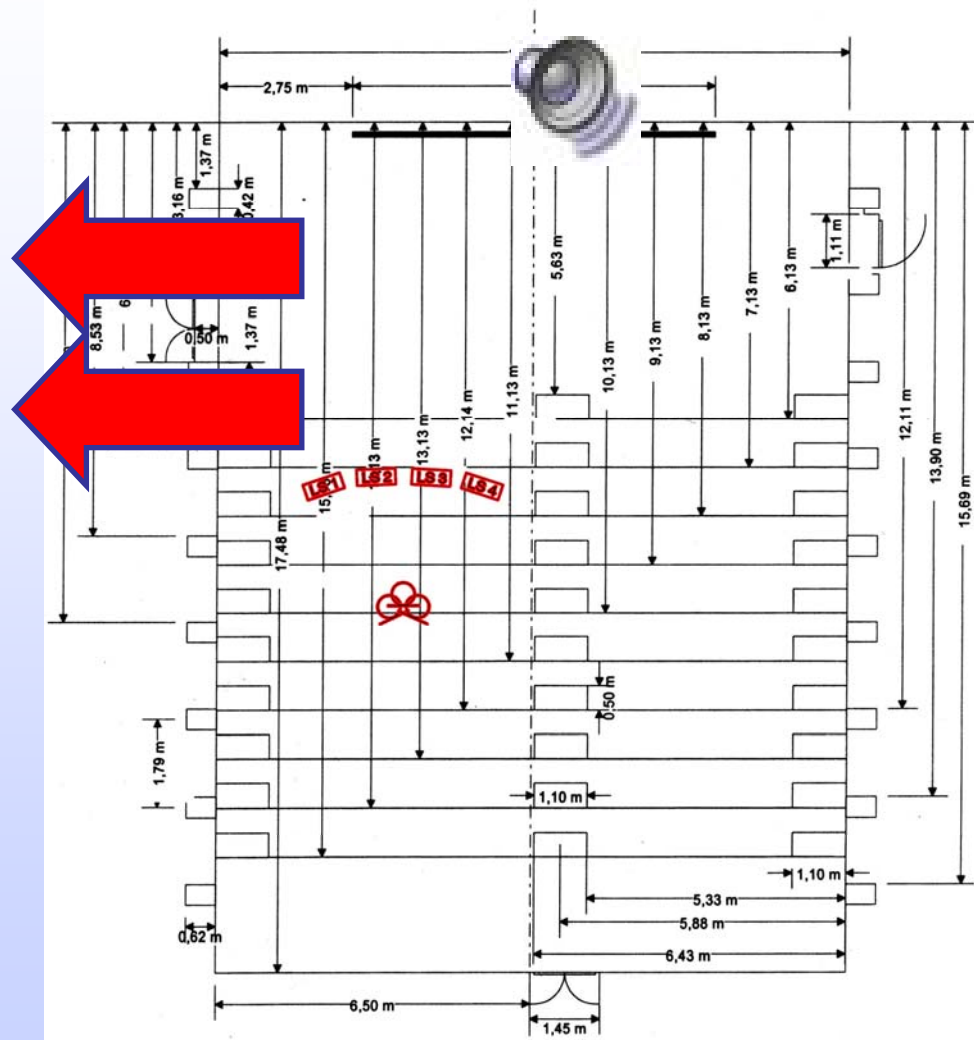
## Anordnung Pos. 2

17

- *DeccaTree*  
*Mit Hamasaki Square*

18

- *INA*  
*Mit quasi ORTF 2*

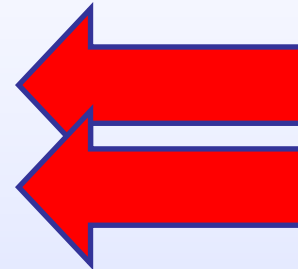




## IRT Kreuz – Sprache/Klavier

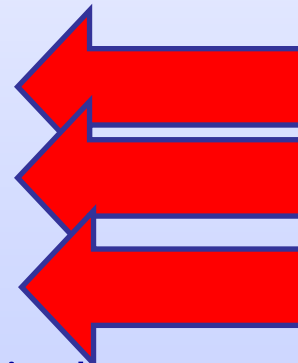
- IRT Kreuz groß

- Surround
- ITU Downmix



- IRT Kreuz klein

- Surround
- ITU Downmix
- ITU + 90 Grad



Phasendreher zwischen  
vorne und hinten

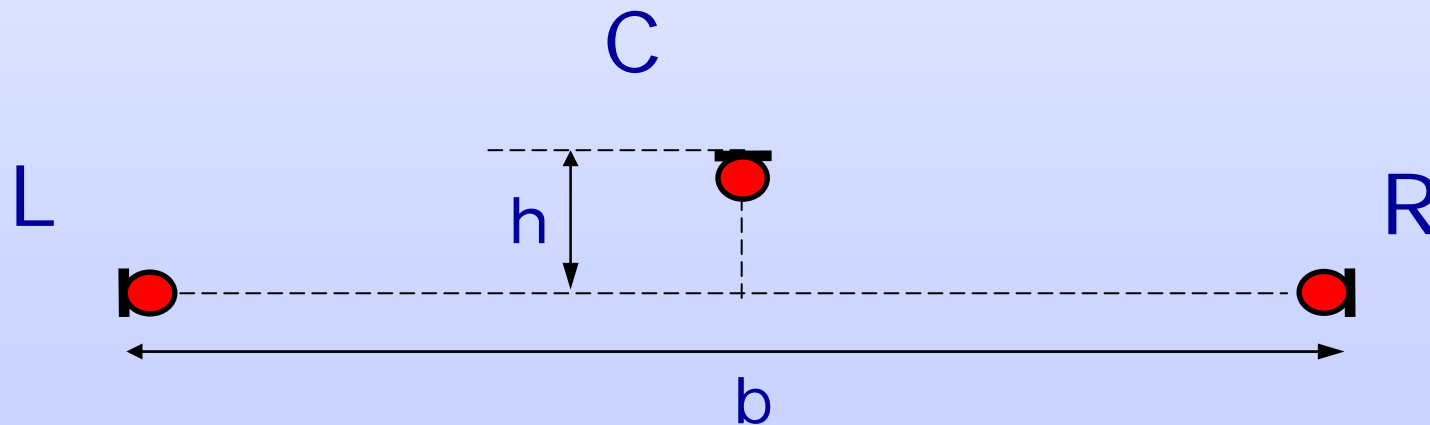
No.  
Spr. Kl.

|                  |    |    |
|------------------|----|----|
| G<br>s-d         | 19 | 24 |
| k<br>s-d         | 20 | 25 |
| G-k<br>d         | 23 | 26 |
| k<br>d-pd<br>+VV | 22 | 27 |



## OCT1

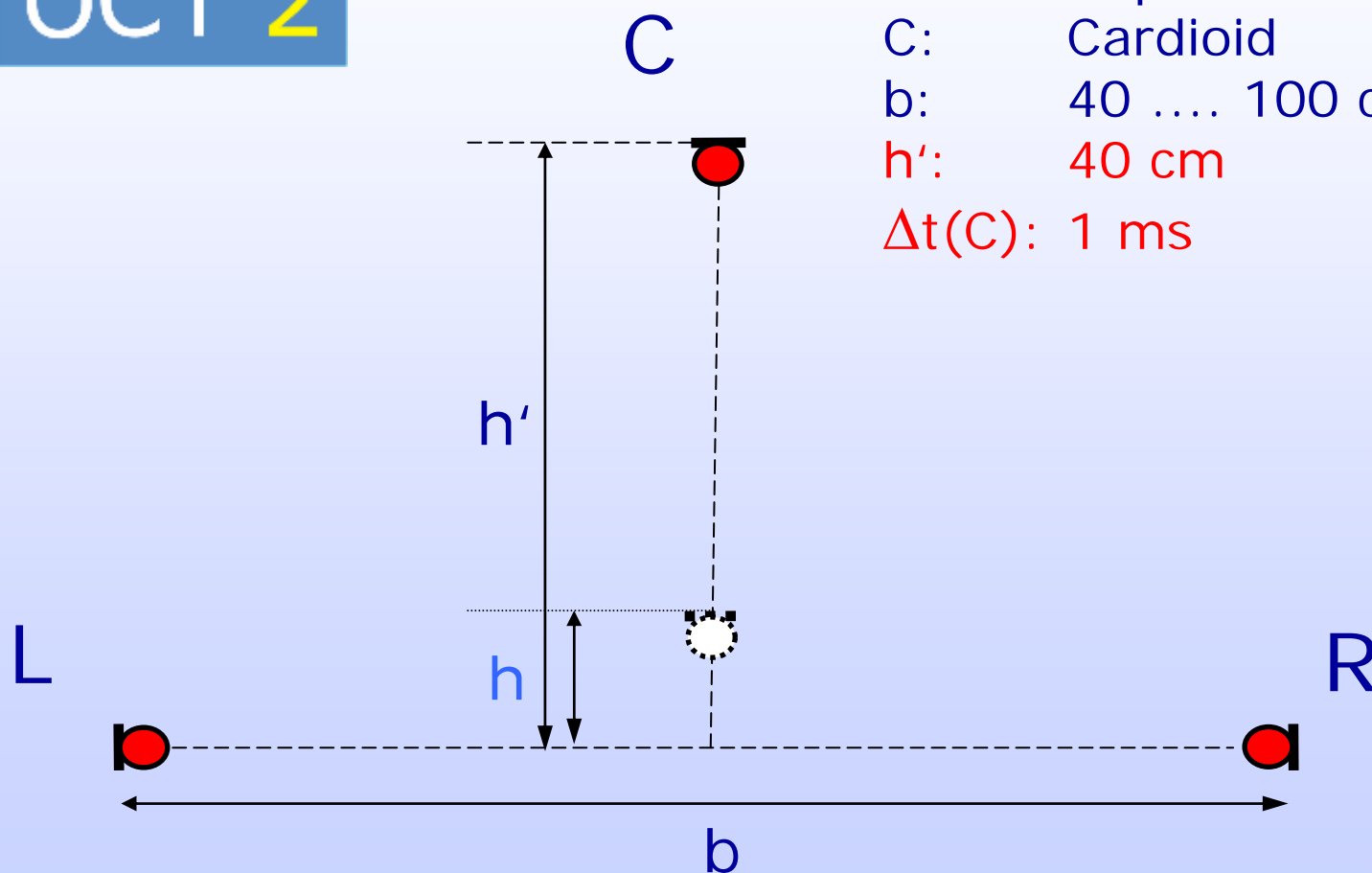
L, R: Super-Cardioid  
C: Cardioid  
b: 40 .... 100 cm  
h: 8 cm





OCT 2

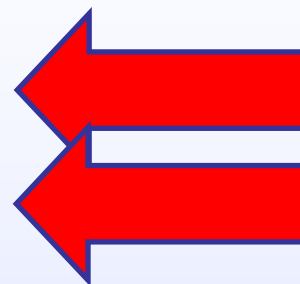
- L, R: Super-Cardioid
- C: Cardioid
- b: 40 .... 100 cm
- h': 40 cm
- $\Delta t(C)$ : 1 ms





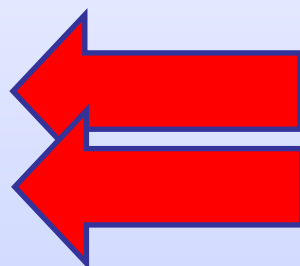
## Surround (L,C,R)

- *OCT1*
- *OCT2*



## Downmix

- OCT1
- OCT2



No.  
E. V.

*G-k*  
*s* 50 46

*G-k*  
*d* 51 48



No.

40

41

Kugel-  
ab



C1,C2

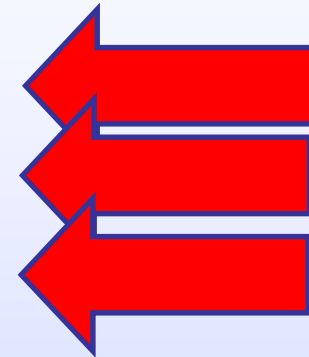
L-R

LS-RS



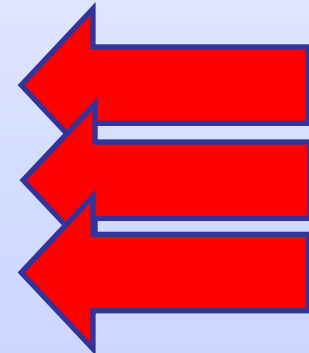
Surround:

- ab
- *OCT1*
- *OCT2*



ITU:

- ab
- *OCT1*
- *OCT2*





Aufbau:

Das vorgeschlagene Hauptmikrofon OCT2 hat im Gegensatz zum OCT ein nach vorne verschobenes Center-Mikrofon ( $h=40\text{cm}$ ).

Diese Verschiebung wird ausgeglichen durch eine elektronische Verzögerung um  $\Delta t = 1 \text{ ms}$ .



### Eigenschaften des OCT2:

- Optimierte Downmix-Eigenschaften im Vergleich zu OCT1 bezüglich **Klangfarbe** und **räumlicher Abbildung** (geringere Korrelation im Nachhall)
- Die präzise und stabile **Richtungsabbildung** des OCT1 bleibt erhalten
- Die Abbildungseigenschaften des OCT1-Surround bezüglich **Tiefe** und **Raumeindruck** bleiben erhalten, das Klangbild wird dem Decca-Tree ähnlicher

→ Mehr Informationen und Klangbeispiele auf

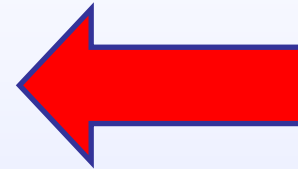
[www.hauptmikrofon.de/oct2.htm](http://www.hauptmikrofon.de/oct2.htm)



DeccaTree (mit AB4)

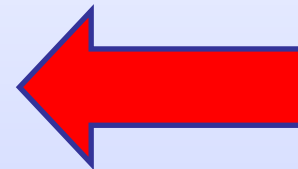
No.

*Surround (L,C,R, Ls, Rs)*



31

ITU Downmix



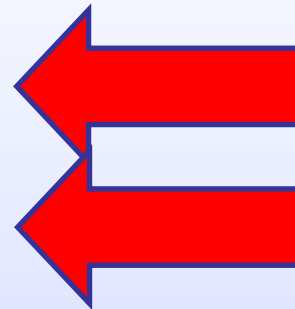


## Decca-Tree - IRT Kreuz groß (Referenz)

No.

### Surround:

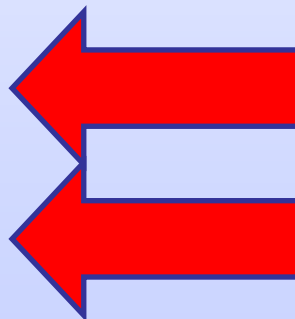
- IRT groß (L,R,Ls,Rs)
- *DeccaTree*



32

### Downmix:

- IRT groß
- *DeccaTree*



33



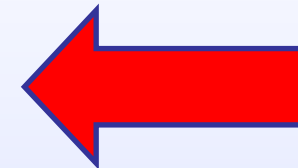
Orts- und Klangfarbenunterschiede an verschiedenen  
Positionen

No.

34

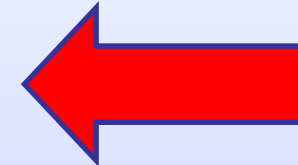
**2.0-Stereo:**

ORTF (Quelle 1→2→3→4→3→2→1)



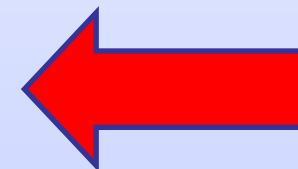
**5.1:**

*Decca (Quelle 1→2→3→4)*



**2.0-Downmix:**

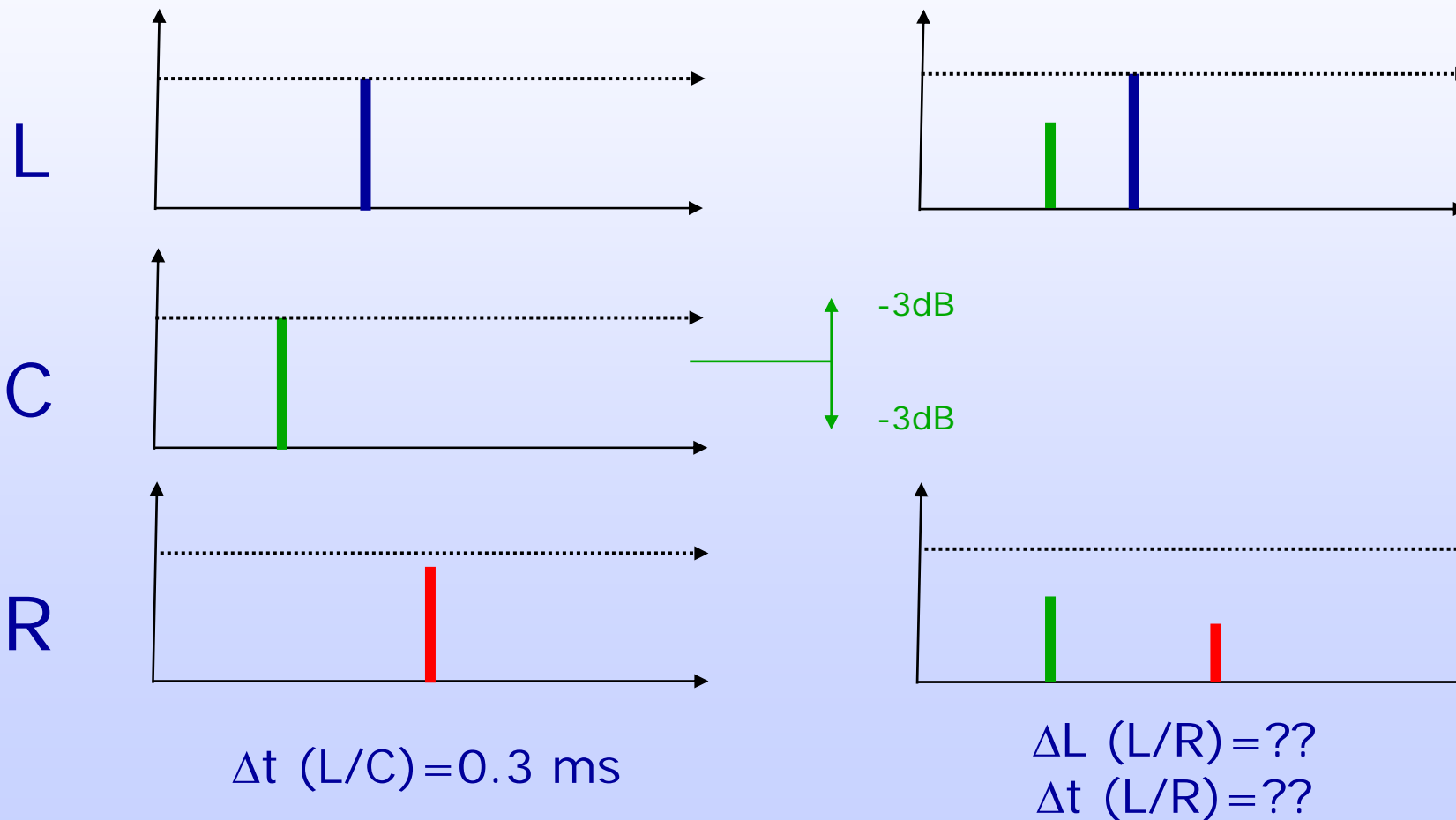
Decca (Quelle 1→2→3→4→3→2→1)





## 5.1

## Downmix





## Kleine Nierenabstände INA – OCT2

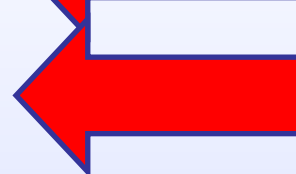
No.

*Surround* INA

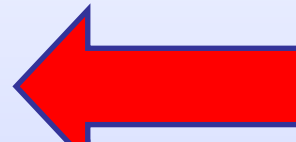


35

*Surround* OCT2

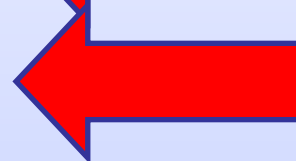


Downmix INA



36

Downmix OCT2



Nachteilig für Downmix:

- Crosstalk über L-C-R(-LS-RS) (Kammfilter + Pegelprobleme)
- Laufzeitvorsprung im Center → [Image Assistant](#)



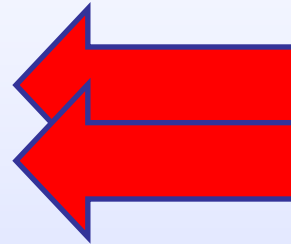
Dekorrelierte, eigenständige  
Signale bei tacet und Sat1

No.

Tacet „Peter & der Wolf“

37

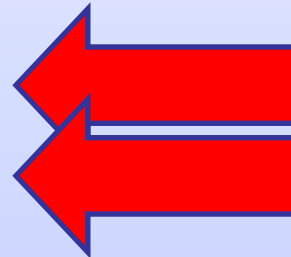
Surround 5.1  
ITU Downmix



Sat1 Champions League

38

Surround  
ITU Downmix





- Wichtige Parameter:
  - Pegelkonstanz, Abbildungsverzerrungen
  - Laufzeit <---> Pegel
  - Crosstalk LCR
  - Korrelation L/LS + R/RS bzw. L/C/R
  - seitliche Reflektionen bei 5.1 erwünscht, bei Downmix aufgrund starker korr. Anteil schädlich!
- Konsequenzen:
  - „Downmixkompatibilität“
  - Wahl der Mittel auch mit Blick auf Downmix-Qualität (Phantasie)
  - Mikrofonierung bzw. Mischung: mit Bedacht wählen
    - Eher größere Abstände → Korrelation verringern
    - Seitliche Reflektionen eher vermeiden
    - LCR-Abstände erhöhen (OCT2)
    - LCR-Crosstalk vermeiden
  - Dekorrelation durch 90° Phasendreher (vorhanden)
  - Parameter für aktive Systeme nutzen



- alle Klangbeispiele auf [www.hauptmikrofon.de/downmix.htm](http://www.hauptmikrofon.de/downmix.htm)
- Danke.