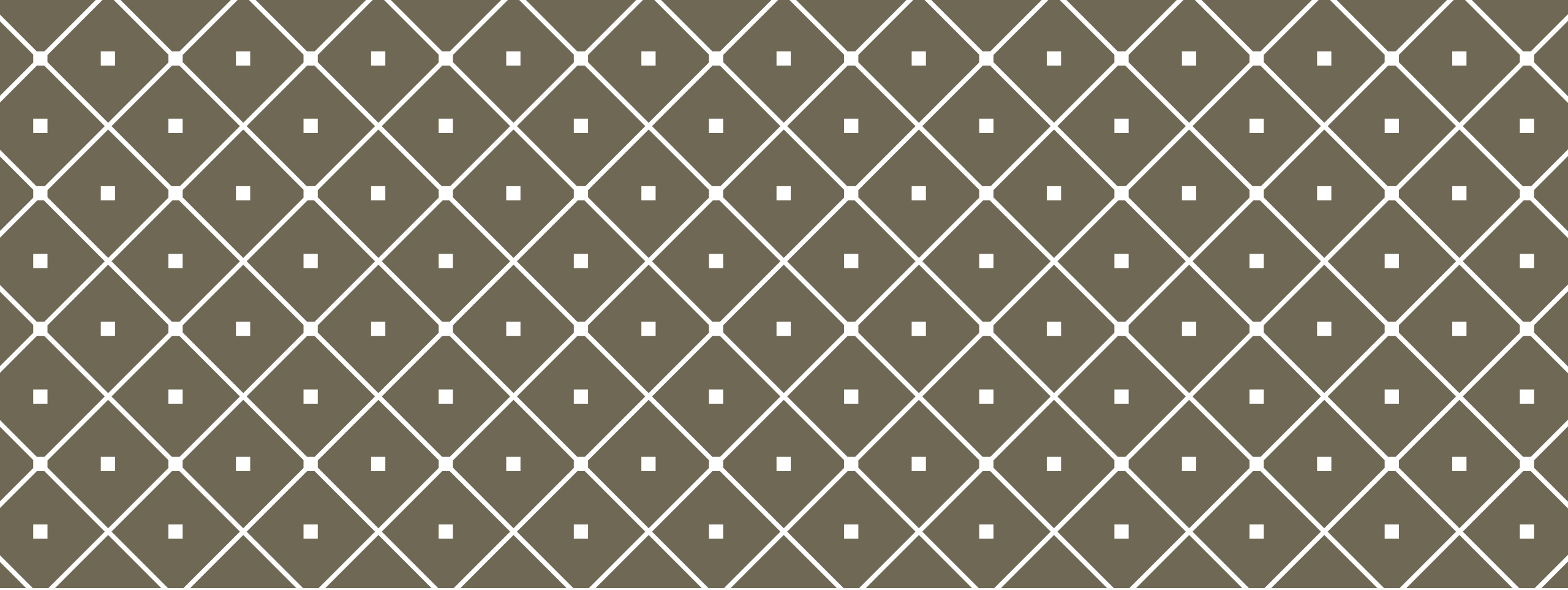


DOCUMENTARY LINGUISTICS I

Prof. **Tomasz Wicherkiewicz**

UAM winter semester 2017/2018

Eighth lecture

28 November 2017

- recording environment
- practical hints
- elicitation and wordlists
- transcribing

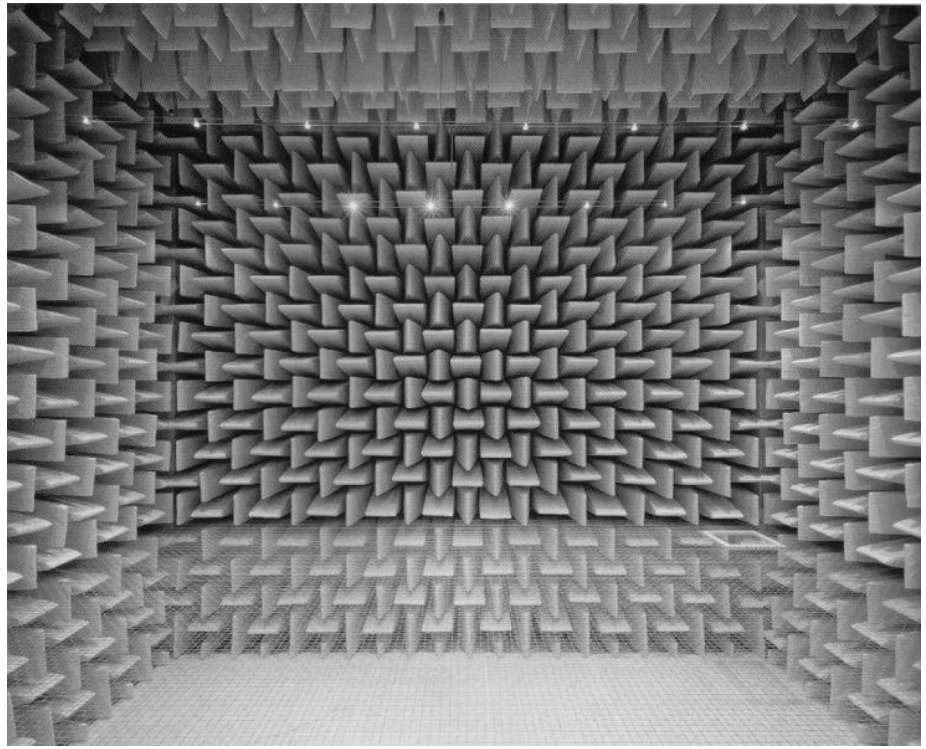
http://languagesindanger.eu/book-of-knowledge/language-documentation

language documentation assumes not only a description of particular elements of grammar and vocabulary, but its central issue is documenting the language in its natural environment, including the characteristics of the speakers, their mutual relationships and the situation in which they live.

When looking at the technical quality we must yet admit that the best recordings can be obtained in an **anechoic chamber** of a recording studio rather than in the language's natural environment.

Noise can be controlled and minimized in the studio, and we may precisely adjust the types and positions of microphones or video cameras in advance, so that even the slightest details will be appropriately recorded.

Such recordings are especially valuable for subtle analyses of the sound system of a language, and that is why anechoic chambers are usually situated in departments of phonetics at universities and research laboratories



Phonetics Labs in the World

EUROPE

- Oxford University Phonetics Laboratory (University of Oxford)
- The Centre for Speech Technology Research (University of Edinburgh)
- University College London Phonetics & Linguistics
- Language and Lingustics (University of Aberdeen)
- Leeds Department of Linguistics & Phonetics (University of Leeds)
- Phonetics Laboratory (University of Cambridge)
- School of Linguistics and Applied Language Studies (University of Reading)
- School of Lingustics, Speech and Communication Sciences (Trinity College Dublin)
- Parlaritaliano (University of Salerno)
- Natural Language Research Group (University of Naples)
- Stockholm Phonetics Lab (University of Stockholm)
- Lund Dept. of Language and Linguistics, Phonetics
- Helsinki Lab of Acoustics and Audio Signal Processing
- Department of Speech Sciences (University of Helsinki)
- Institute of Phonetic Sciences, Amsterdam
- Phonetics Laboratory Leiden
- Laboratory of Phonetics (Paris VII University)
- Laboratorie Parole et Langage (University of Provence, Aix-en-Provence)
- Institute of Phonetics and Digital Speech Processing (University of Kiel)
- Institute of Phonetics and Speech Processing (University of Munich)
- Institute of Phonetics (Saarland Univeristy)
- Experimental Phonetics Group (University of Stuttgart)
- Phonetics Laboratory (Aristotle University of Thessaloniki)

UNITED STATES

- UCLA Phonetics Lab
- Haskins Laboratories (University of Yale)
- Phonetics and Experimental Phonology Laboratory (New York University)
- OSU Phonetics Teaching Lab (Ohio State University)
- Berkeley Phonology Lab (University of California)
- Phonetics Laboratory (Univerisity of Victoria)
- USC Phonetics Lab (University of Southern California)
- UW Phonetics Laboratory (University of Washington)
- Arizona Phonological Imaging Lab
- Phonetics Laboratory (University of Pennsylvania)
- Douglass Phonetics Laboratory (University of Arizona)
- Cornell Phonetics Laboratory (Cornell University)
- Laboratory of Phonetics and Psycholinguistics (University of Kansas)

CANADA

- Laboratoire de sciences phonetiques (University of Montreal)
- Alberta Phonetics Laboratory (University of Alberta)

CHINA

- Phonetics Laboratory (City University of Hong Kong)

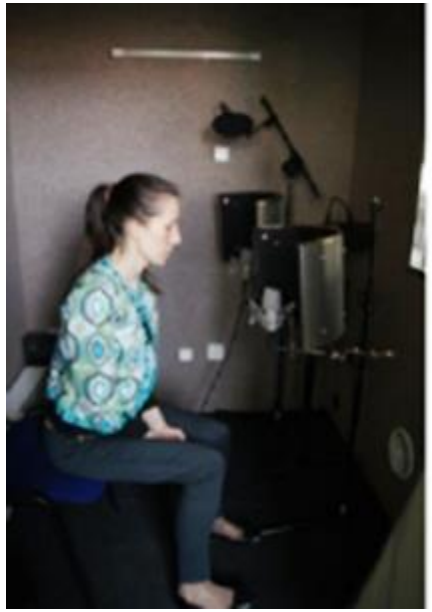
AUSTRALIA

- Phonetics Laboratory (University of Melbourne)

Speech Labs all over the World

- Laboratory of Speech Acoustics of Department of Telecommunications and Media Informatics a BUTE
- Speech Technology Laboratory
- CIRASS, Italy
- FUB, Italy
- LOQUENDO, Speech Technologies & Services, Torino, Italy
- ISTC, Instituto Di Scienze E Technologie Della Cognizione Sezione di Padova "Fonetica e Dialettologia" Consiglio Nazionale delle Ricerche, Padova, Italy
- Speech Information Processing Laboratory, Department of Computer Science, Korea University
- Speech technology group, Department of Electronics and Telecommunications, NTNU, Norway
- DLP, Lund Univ., Sweden
- KTH, Speech Communication Dept., Stockholm, Sweden
- Speech Processing Group, TIK/ETH, Zurich, Switzerland
- LS Dept. Language and Speech, Kat. Univ. Nijmegen, The Netherlands
- SVR, Cambridge Univ., UK
- Speech and Hearing Group, Dept. of Computer Science, University of Sheffield, UK
- ASEL Appl. Sc. & Engin. Labs, Speech Lab, Philadelphia, PA, USA
- AT&T Labs Research, Florham Park, NJ, USA
- CSLU Center for Spoken Language Understanding, Univ.Boulder, Colorado, USA
- Lucent Tech. Bell Labs, Murray Hill, NJ, USA
- MicroSoft SRG Spee[USA]chnology, MicroSoft Research
- STL - Speech Technology Laboratory, Santa Barbara, CA, USA
- SAIL - Speech Analysis and Interpretation Laboratory, USC Viterbi School of Engineering , Los Angeles, USA

- CSLAB Computer Sciences Lab., Australian National Univ., Canberra, Australia
- MARCS, Auditory Laboratories, Sydney, Australia
- STRG Speech Technology Research Group, Univ. Sydney, Australia
- Speech and Language Processing (University of Macquarie) Australia
- Signal Processing and Speech Communication Lab, Graz University of Technology, Graz, Austria
- ESAT, Speech Group, Univ. Leuven, Belgium
- Speech Prosody Studies Group, University of Campinas , Brazil
- LAFAPE - Laboratório de Fonética Acústica e Psicolinguística Experimental, Campinas, Brazil
- ERMETIS Equipe de Recherche en MicroElectronique & Traitement Informatique des Signaux, Univ. Quebec, Chicoutimi, Canada
- National Centre for Audiology, Canada
- Speech Processing and Transmission Laboratory (Laboratorio de Procesamiento y Transmision de Voz)
- Center for Speech Technology Tsinghua University
- Speech@FIT group, Faculty of Information Technology, Brno University of Technology, Czech Republic
- CTU Speech Group, Czech Technical University, Prague, Czech Republic
- CPK - Center for PersonKommunikation, Aalborg, Denmark
- Speech and Image Processing Unit (SIPU), Joensuu, Finland
- Departement Parole - Cognition de GIPSA-Lab (ex. ICP) - Inst. of Spoken Communication, Grenoble, France
- IRIT - Institut de Recherche en Informatique de Toulouse, France
- Laboratoire Parole et Langage, Univ. de Provence, France
- LIA (Laboratoire Informatique d'Avignon), Avignon, France
- LIMSI-CNRS, Orsay, France
- BAS - Bavarian Archive for Speech Signals, Inst. Phonetics, University of Munich, Germany
- NatS Natural Language Systems Division, Computer Science Dept., University of Hamburg, Germany



Studio recording environment:

recordings in an anechoic chamber,

the Laboratory team at work,

equipment (the Laboratory of the Psycholinguistics Department, Adam Mickiewicz University in Poznań,

photos: Agnieszka Czoska & Maciej Karpiński

compromise between quality control and natural environment requirements

it is practically impossible to capture most real communicative events in an artificial surrounding of a studio.

This may be particularly true for elderly speakers who sometimes are the only speakers left of a severely endangered language.

On the other hand, recording speech outside the studio usually means difficulties in achieving good quality, even if we use excellent recording devices.

When we sit in a room and chat, we usually don't pay attention to small background noises, but when we listen to a recording of that conversation we suddenly discover that there was a clock ticking or a fridge buzzing!

Observer's Paradox

Speakers who are recorded (even in a very friendly environment) usually pay more attention to their way of speaking and consequently modify their linguistic behaviour: their speech can change in a quite unpredicted way.

It can become more/less formal, more/less polite or politically correct which can be manifested by changes in phonetic-acoustic parameters such tempo, intonation, intensity, timing patterns, pausing schemes, and others.

Project Team

- helping the linguist to learn the language
- recording, transcribing, and translating
- editorial work
- helping the linguist to understand and translate the recordings
- dictionary work

Mosel 2016
Fieldwork and community language work

Elicitation

- reliable, especially in the very beginning of fieldwork

Elicitation means getting linguistic data from native speakers by asking questions.

some older fieldwork manuals give advice on what kind of questions to ask or not to ask, how to make the interview interesting and keep the informant attentive, etc.

In this manner, such manuals quite automatically assign a passive role to the native speaker.

If fieldwork is to be a mutual teaching-learning event, this approach is no longer acceptable.

Rather, methods to develop in order to involve the speaker as an active partner who eventually becomes an independent language documenter.

Wordlists

to investigate the phonological system and create a working orthography or understand an existing orthography

In traditional fieldwork manuals - compiling wordlists by asking bilingual native speakers for the translation of wordlists in the lingua franca into their native language

Method questionable on both linguistic and psychological grounds.

The native speakers might feel embarrassed when asked for the translation of a word they do not understand or even worse, a word that they cannot translate because they have forgottenthe indigenous equivalent, or because there is a taboo about it.

Alternative method / Mosel 2006

– explain what you need the wordlists for – this is not just for studying phonology and orthography; the first wordlist of about 180 words will also serve as the starting point to build short clauses;

– discuss what semantic fields might be suitable to start with, and perhaps suggest food and cooking;

– ask the native speaker to teach you words of this particular semantic field by dividing it into subcategories, e.g.:

– fruit and vegetables, edible animals

– dishes

– activities

– tools.

Thus you ask:

– tell me the names of fruit and vegetables you grow and eat

– what do you do when you make a dish with potatoes?

– what kind of things do you use?

Transcription

If the local language workers are literate in any language, they can be asked to make transcriptions. Even if their spelling is inconsistent or neglects important distinctions (ones that linguists might consider indispensable), their transcriptions will be a great help.

The most important thing to teach them is to transcribe what the speaker actually says and not to correct speech errors and other mistakes, although such editing is certainly legitimate in later stages of data collection and analysis.

In order to allow for a genuine participation of the speech community in the documentation project, it is imperative that all recordings are transcribed in a practical orthography readily accessible to literate but not linguistically-trained native speakers.

For specialists interested in phonetics and phonology, only a selected corpus needs to be rendered in a phonetic transcription.

The more time spent on narrow phonetic transcriptions, the smaller and the less useful the corpus of annotated recordings will be for the speech community and for researchers who are not interested in phonetics and phonology

The local language workers may be afraid of “spelling mistakes” in their work. But as long as the orthography has not been standardized, there is no such thing as a right or a wrong spelling, and they should be encouraged to follow their intuitions, which may be relevant for the analysis of the phoneme system.

As discussions on spelling problems and standardization can be quite emotional and are often guided by sociopolitical issues, they should be postponed to a later stage when the linguist is more familiar with the speech community and the local language workers have gained more experience in writing their language.

However, for the data base of the project, especially for the lexicon, a consistent working orthography that distinguishes between norms and variants is a prerequisite, but this does not necessarily imply that the local transcribers have to learn and use it.

Later, when the speech community decides on their own norms, the working orthography can be adjusted to their standard orthography.

Alternative method + / Wicherkiewicz

Józef Gara

Bibowa w ogrodzie

Stroje wilamowskie

http://inne-jezyki.amu.edu.pl/Frontend/Language/Details/10

dysk

The Basic Principles of Creating Audiovisual Material

Marta Ostajewska, Artes Liberales, UW

© !

Check audiovisual equipment

- Make a list of the things that you need and check if you have everything
- Do you have spare batteries and memory cards?
- Does the camera and the microphone work?
- Test an equipment by making a trial recording
- Check audio during trial recording: is the microphone firmly taped and positioned on the right channel?

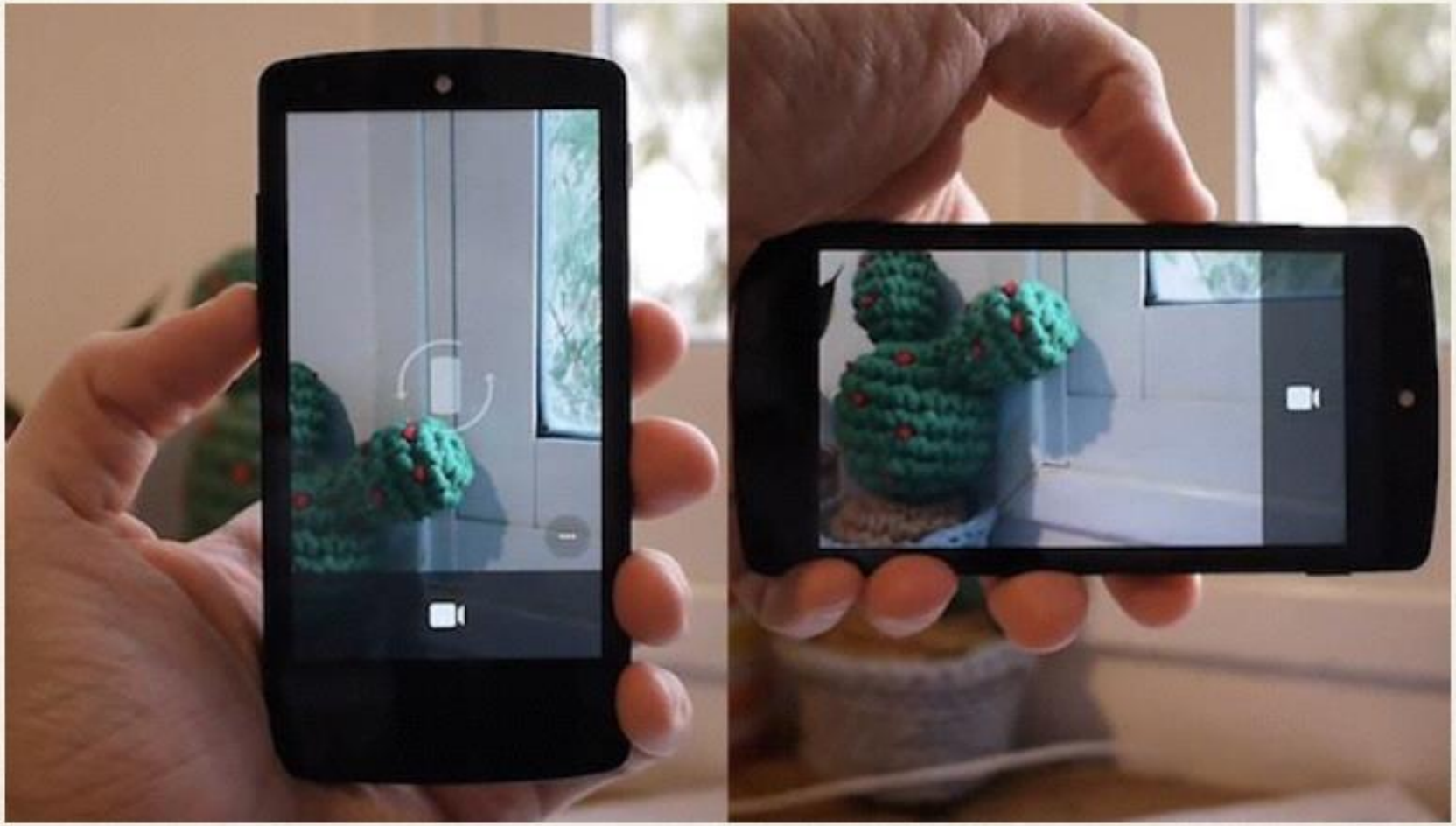
Common mistakes

The material is not recorded

Always check whether the recording button is pressed, whether the battery is charged and if you have a spare memory card

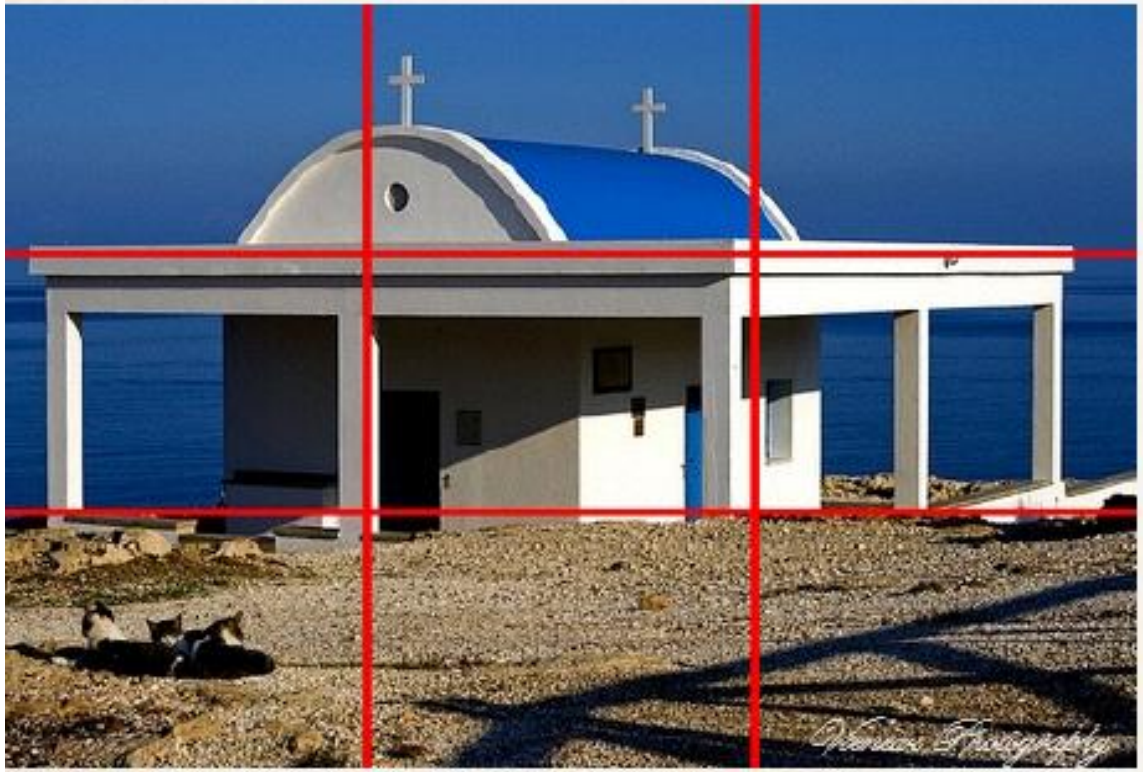
Rotation from horizontal level to vertical level

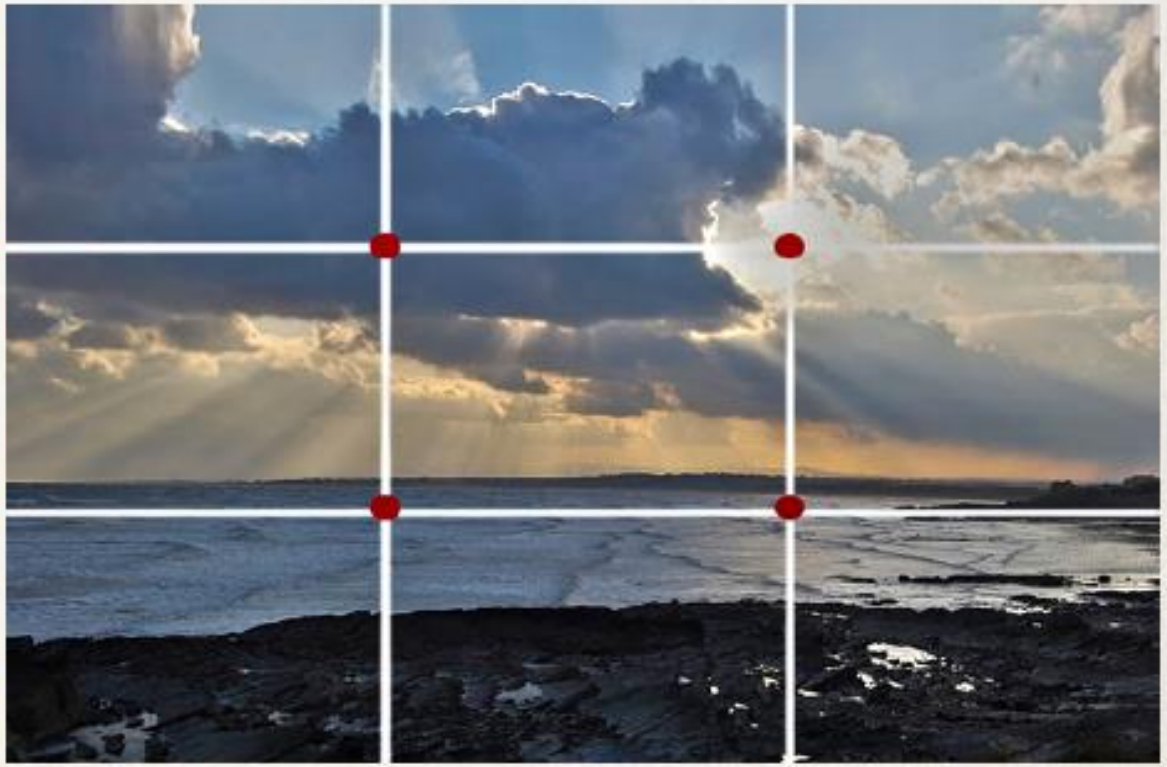
Do not turn the camera, shoot horizontally, it makes your montage easier



Crooked images

Look at the lines and try to keep the camera even, level to the horizontal lines – frame of the buildings. Always level a tripod before recording.





Shortcuts of the perspective

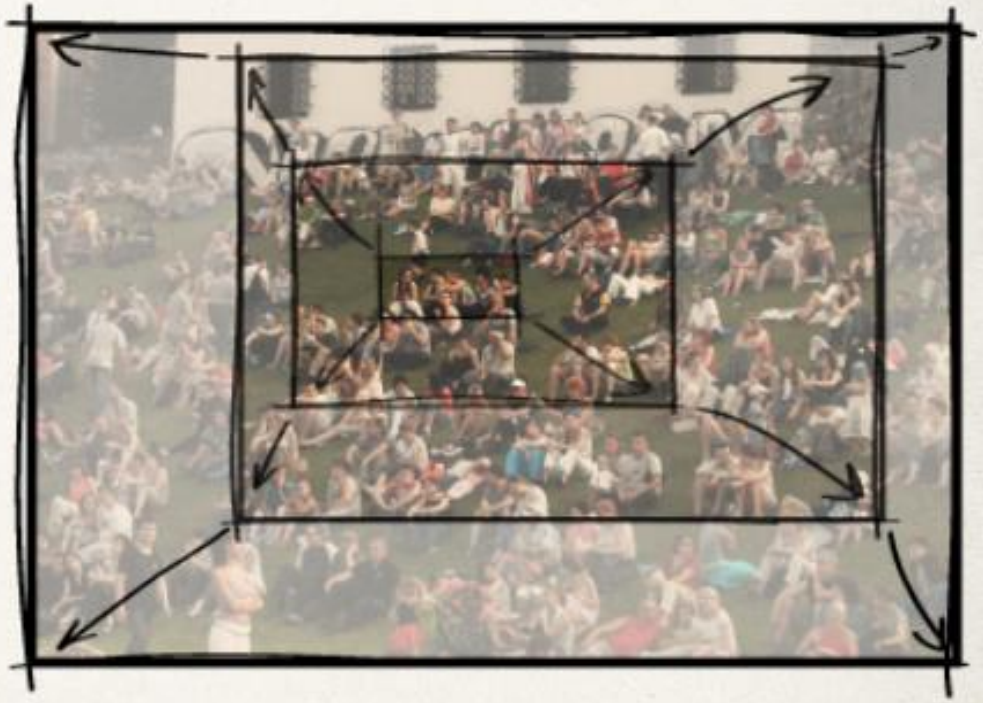
Do not raise up the camera. Move as far away from the building as possible and zoom in (if you have the opportunity) or – preferably – stand on some elevation (hill, bench, stairs, etc.).

Bad lighting

- Do not film a caller under the sun and against the window, in a place with variable light. Shot a scene preferably on a neutral background
- Record under daylight in a well lit room
- Do not mix artificial light with natural light

Moving to much

- Give a scene time to resound (do not change the position of the camera constantly). Try to be calm during filming
- **Do not make jumpy zoom in – zoom out**
- Get closer to the subject with the camera instead of zooming
- If you zoom in with a camera on the tripod, set a good frame and give the scene a time



Out of focus

Turn on manual focus, zoom in, sharp the image, zoom out

Positives of autofocus

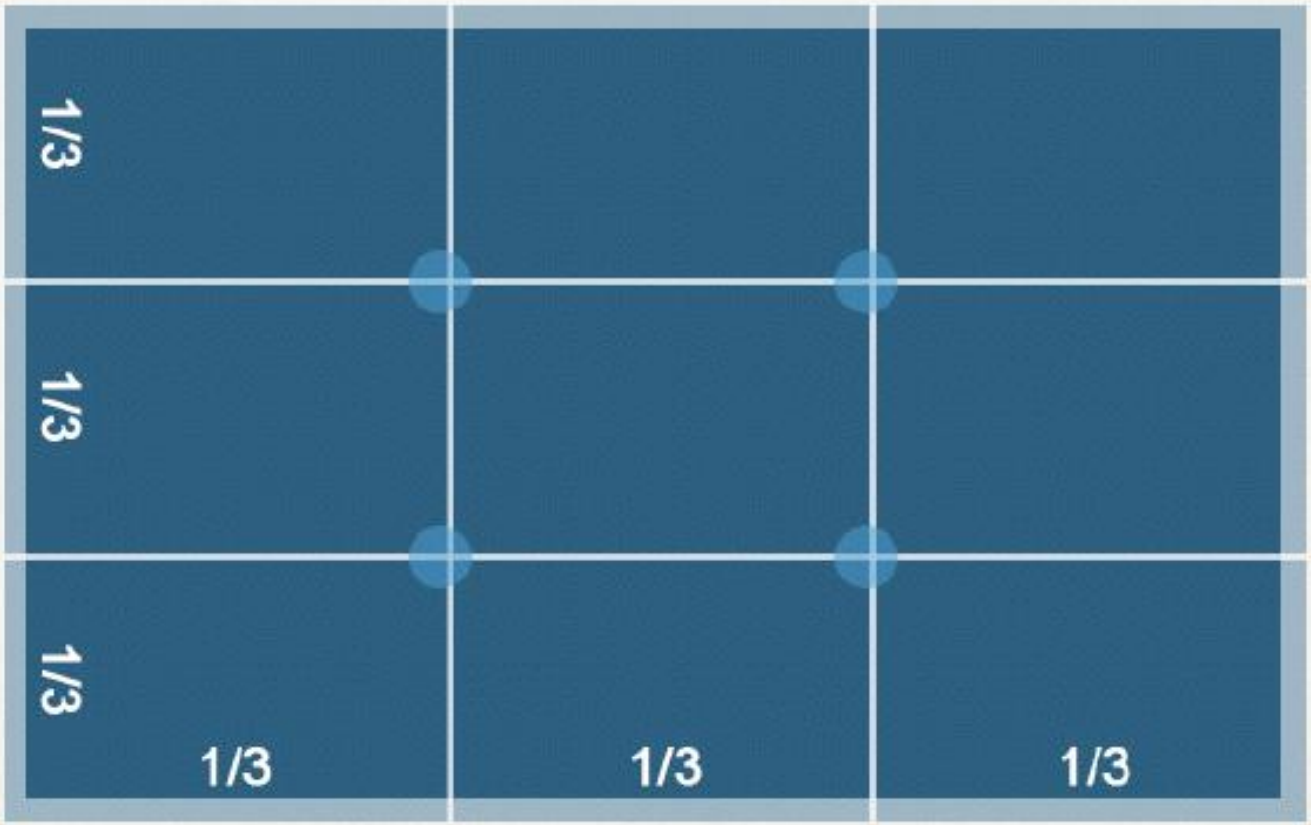
- For beginners
- It works in incalculable conditions
- It sharpens on a particular element from the frame
- It allows to focus on other aspects of film - composition, movement, etc.

Negatives of autofocus

- Loss of focus when moving in the frame
- Lack of full control over depth of field
- Problem with changing lighting, blurring focus
- Engine noise

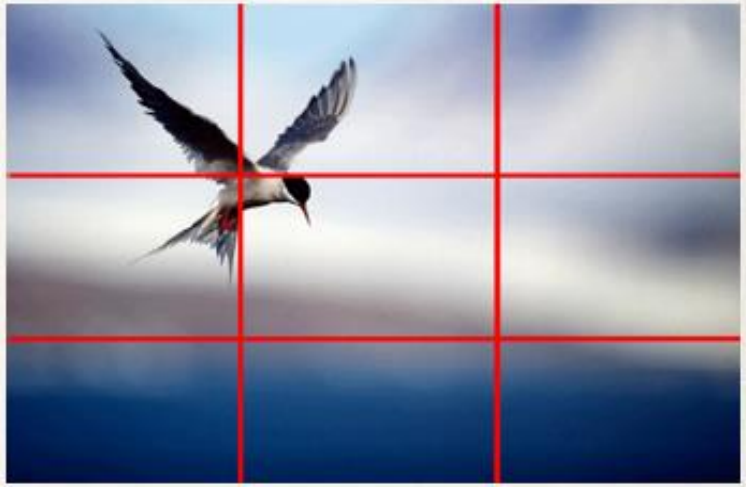
The composition of the picture

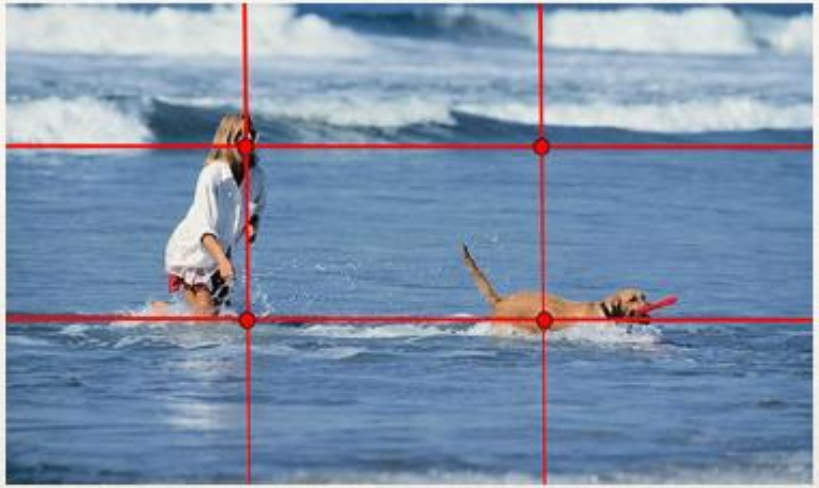
Lookroom, Headroom, The Rule of Third



The Rule of Thirds

An image should be imagined as divided into nine equal parts by two equally spaced horizontal lines and two equally spaced vertical lines, and that important compositional elements should be placed along these lines or their intersections

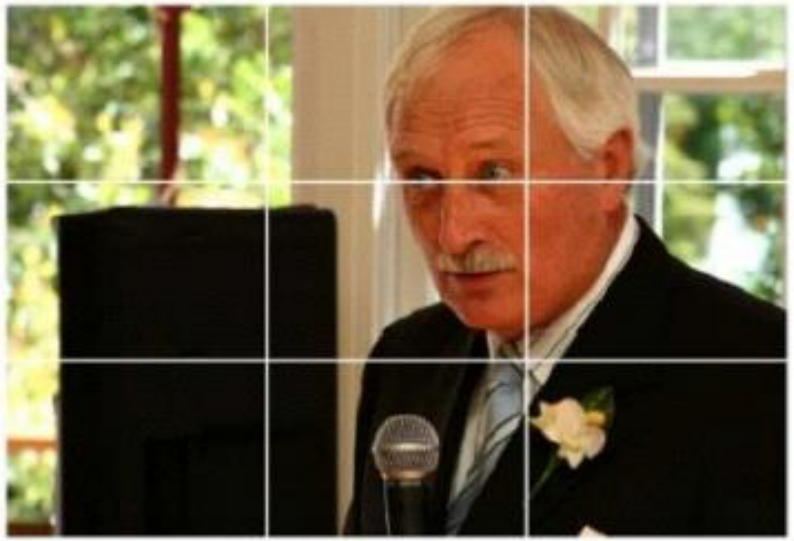




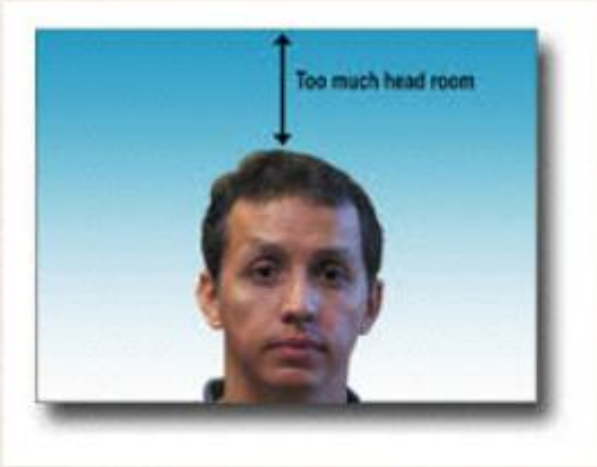
The Rule of Thirds

The Rule of Thirds

When shooting a close-up of the face, the eyes are the center of attention : divide the screen into thirds and try to compose your shot so they're at one of the points where the lines intersect.



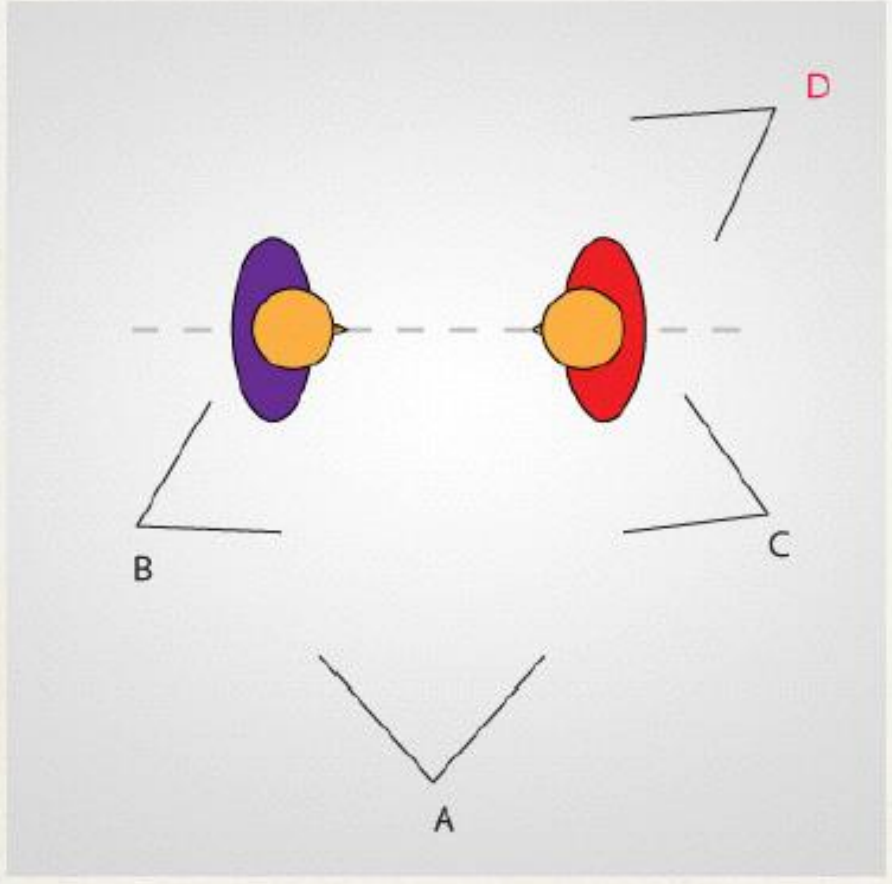
Lookroom and Headroom

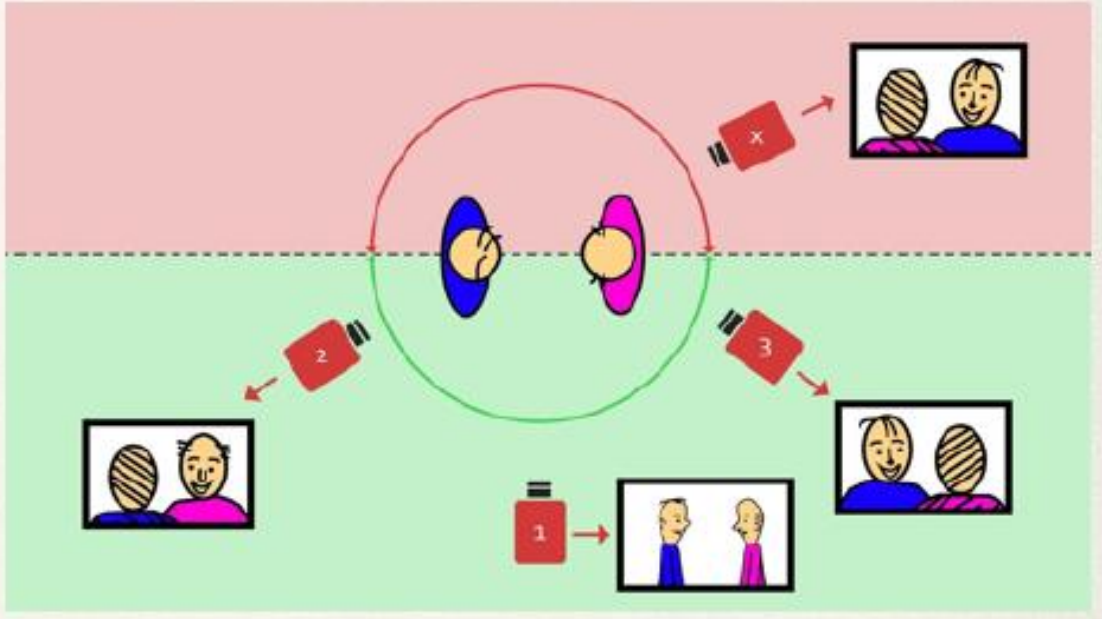


The Lead Room is the space in front of where the actors are looking. In this shot from *Butch Cassidy and the Sundance Kid* (1969), the lead room is on the left side of the frame.

180 degree rule

It allows you to record dialogue between characters correctly. According to this principle, we have to lead an invisible line crossing the two recorded callers and simply do not cross it during recording. (A, B, C – correct, D – incorrect)





Reasons you should use video in language documentation

- Gesture is an important part of communication
- Gesture is an important part of cognition
- Gesture is an important part of culture
- Having video as well as audio makes transcription easier, and ensures that the elements of discourse that are not in the spoken channel are still collected

Video and language documentation

- You have to discuss sharing and permissions with the community, and the individuals you are recording with and have the permission to record
- Setting up any recording situation is obtrusive
- Being comfortable with your equipment and making people feel comfortable with your presence mitigates many of those problems. Practice setting up as many times as you can before you begin the project. Record your friends and family. Know your equipment well enough now to continue chatting throughout the setup.

In video documentation of the language it is important that the gestures fit in the frame, do not go beyond the frame, are not cut

Bad framing: cutting people's legs and heads

- Try to be consistent in framing: if you change plan from far to near, from near to far pay attention to where do you stop framing

- Try not to "cut" the body at the height of the joints (ankles, knees, hips, etc.)

- When you are filming humans use: Wide shot, Medium shot, Medium close up, Close up, Extreme close up

Wide shot

Medium shot

Medium close up

Close up

Extreme close up

Story telling (add three plans)

- If you shot an event, except for the shooting the action itself, make an additional three plans
- Broad plan: A broad plan covering the area
- Changing angle: Place from another angle, e.g. a view of people watching events
- Detail: Closeup on smaller objects, details

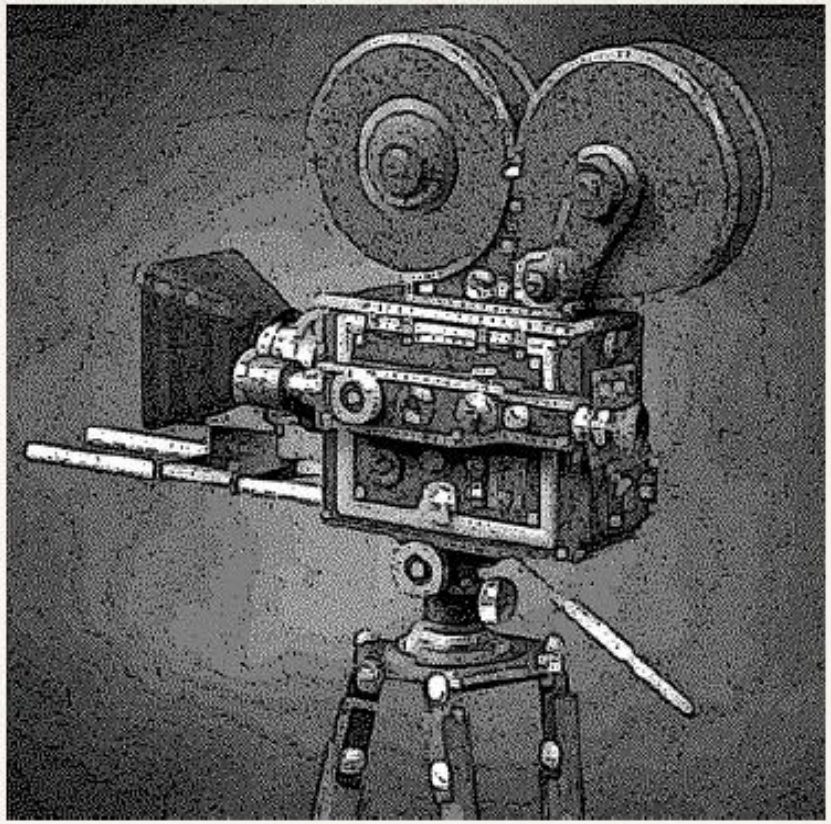

STOP

Inter-cuttable shots

A kind of shot helping to smoothly integrate scenes that do not take part in the narrative mainstream. For example, "a talking actor," "a ticking clock (inter-cuttable shots)," "a talking actress".

Sound problems

- Do not record sound only on the camera, have an external microphone
- Try not to record in the strong wind
- Dress up your microphone outdoor
- Avoid overdrive




ENJOY!

Wordlist Elicitation in Hałcnów a Very Small &/or Extremely Endangered (Micro)Language