Laurence Nigay
Laurence.Nigay@univ-grenoble-alpes.fr

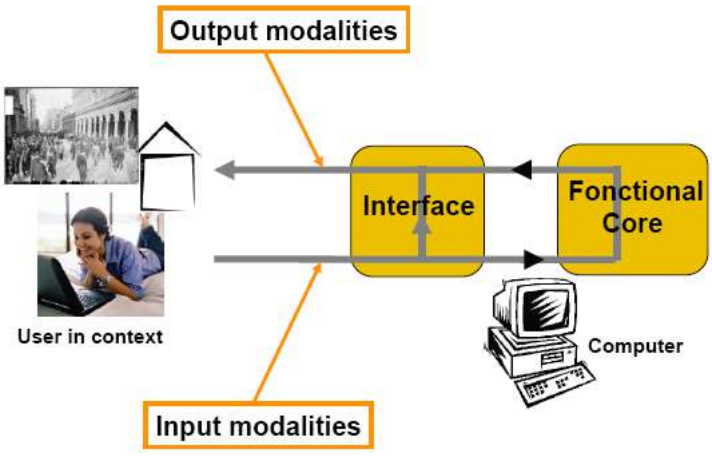
Multimodality: design

Underlying concepts
Design space
Rules of thumb, heuristics

1

1

Underlying concepts



2

2

The Pipeline Model

- 2 concepts as point of contact between the user and the system:
  - interaction language
  - physical device
- Interaction language: set of well formed expressions used by the system or the user to exchange information
- Inter. language & phys. device = 2 facettes of an expression
  - interaction language = the structure (Hemjslev's form)
  - physical device = the observable (Hemjslev's substance)

3

3

Underlying concepts
Definition of a modality

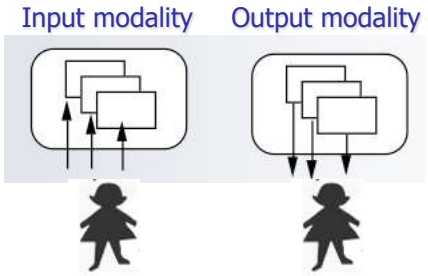
- Built-in cognitive capability of the system for interpretation and rendering
- Input modality
  Interpretation function: sequence of transformations from input "raw information"
- Output modality
  Rendering function: sequence of transformations to output "raw information"

4

4

Definition of a modality

- Modality = (device, interaction language)
  - A set of sensors (input devices) or effectors (output devices)
  - A processing facility based on a language



5

5

Definition of a modality

- Modality = (device, interaction language)
  - A set of sensors (input devices) or effectors (output devices)

Perception/Action

  - A processing facility based on a language

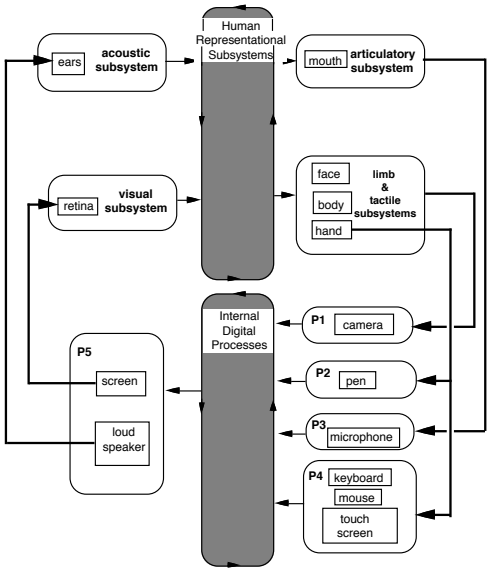
Cognition

6

6

Definition of a modality

- Theory ICS



7

7

Definition of a modality

- Modality = (device, interaction language)
- Multimodality
  - Multi device Mono Language
  - Multi device Multi Language
  - **Mono device Multi Language**
  - e.g. table and graph displayed on screen as two different modalities
    - M1 = (screen, table) and M2 = (screen, graph)

8

8

Definition of a modality

- Modality = (device, interaction language)
- Interaction paradigms such as perceptual User UI tangible UI, embodied UI and AR open **a vast world of possibilities**
  - M1 = (microphone, natural language)
  - M2 = (keyboard, command language)
  - M3 = (mouse, direct manipulation)
  - M4 = (smartphone, 3D gesture) ***embodied UI***
  - M5 = (HMD, 3D graphics) ***AR***
  - M6 = (bottle-sensor, 3D gesture) ***tangible UI***
  - M7 = (GPS, localization) ***perceptual UI***
  - *M8 = (Tongue display, 2D shape)*

9

9

Definition of a modality

- Input Modality = <d, I>

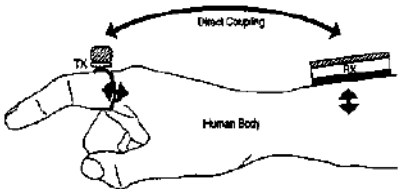
Speech = <  , natural language>

10

10

Definition of a modality

- Input M = <device, text>



11

11

Definition of a modality

Input M =
<camera, Gesture>

Input M =
<microphone, NL>

Input M = <Smartphone, Gesture>
Embodied modality

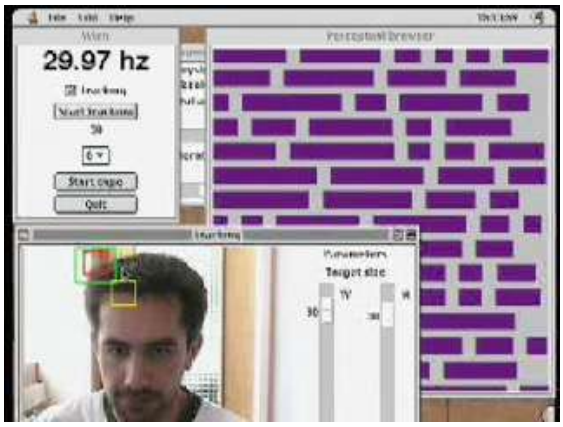
Input M =
<stylus, direct manipulation>

12

12

Definition of a modality

- Input M = <camera-head, gesture>

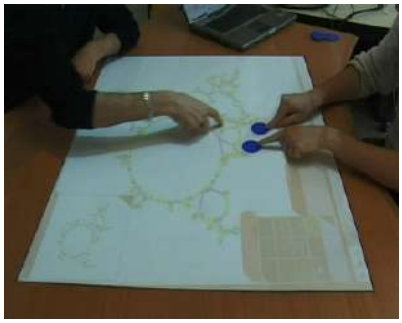


13

13

Definition of a modality

- Input M =
  <camera-token, direct manipulation>



14

14

Definition of a modality

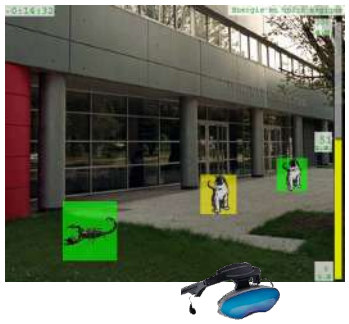
- Input M = <bottle-sensor, gesture>

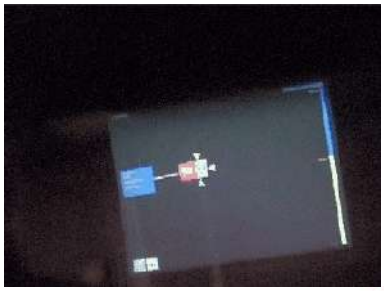
15

15

Definition of a modality

- Input Modalities (*sensing modalities*)
- M1 = <GPS, localization>
- M2= <magnetometer, orientation>





16

16

Definition of a modality

- OUTPUT Modality = <d, l>

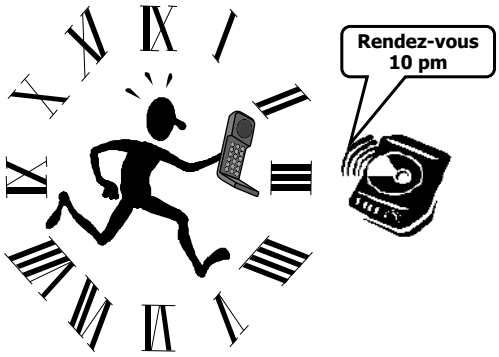
  M = <HMD, 3D graphics>

17

17

Definition of a modality

- Output M = <loudspeakers, NL>
- 3D sound:

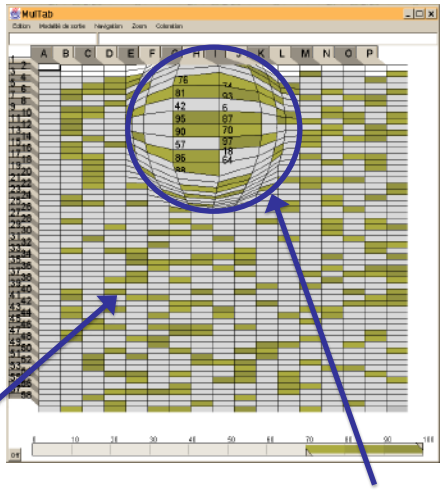


Soundbeam Neckset

18

18

Definition of a modality



Output M1 =
<screen, table>

Output M2 =
<screen, deformed table>

19

19

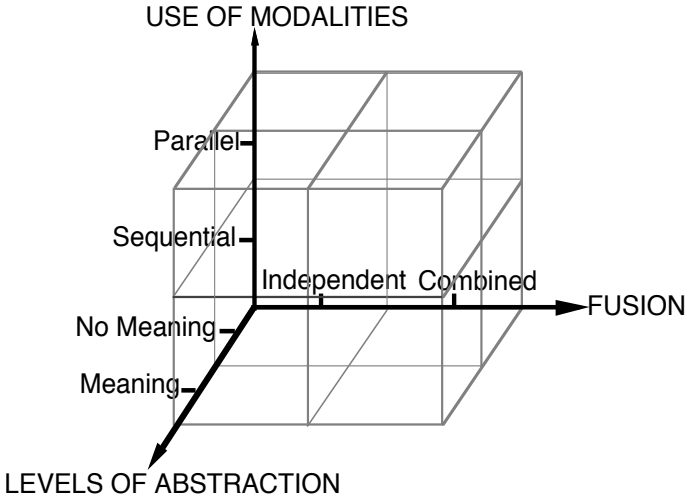
Underlying concepts

- Modality = (device, interaction language)
  - Input modality
    - Interpretation function: sequence of transformations from input "raw information"
  - Output modality
    - Rendering function: sequence of transformations to output "raw information"
- Four intertwined ingredients (for both):
  - 1. Levels of abstraction
  - 2. Context
  - 3. Fusion and fission
  - 4. Granularity of concurrency

20

20

Underlying concepts

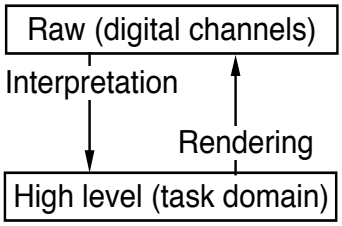


21

21

Dimension 1: Levels of Abstraction

Expresses the variety of representations supported by the system:



Interpretation function: Ability to abstract

Rendering function: Ability to materialize

22

22

Dimension 1: Levels of Abstraction

Example: Speech input and output

Interpretation function **Ability to abstract to**	Rendering function **Ability to materialize from**
Digital signal	Symbolic representation of meaning
Word or a pattern of words	Pre-stored text message (text to speech)
Meaningful sentence	Pre-recorded vocal message

We consider two values only:
MEANING / NO MEANING

23

Dimension 1: Levels of Abstraction

- The capacity of abstraction may vary with the context
- Example : text editor
  - command mode: text is processed -> high level
  - input mode: text is recorded only -> raw
- Context of commands
  high level interpretation
- Context of task-domain data
  low level interpretation

24

Dimension 2: Use of Modalities

- Supported use of modalities
- Sequential:
  Use of the modalities one after another
- Parallel:
  Use of multiple modalities simultaneously
  - Multiple devices used simultaneously

25

25

Dimension 3: Fusion

- Independent: (Absence of fusion)
  Independent interpretation/rendering process for each modality
- Combined: (Presence of fusion)
  Fusion of data expressed using different modalities

26

26

Dimension 3: Fusion

- Combined: Combination of chunks
- It occurs at multiple levels of abstraction

- Lowest level: chunks from distinct modalities
- Higher level: chunks from dictinct contexts

27

27

Dimension 3: Fusion

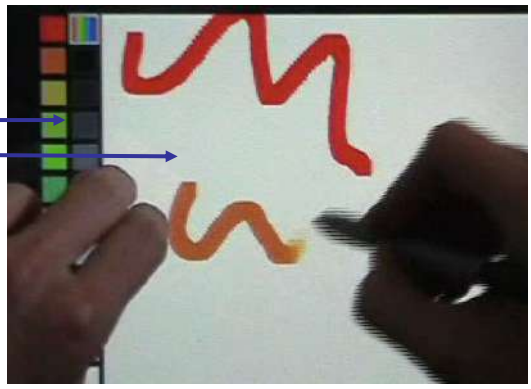
- **Lowest level: chunks from distinct modalities**
- Fusion of data expressed using different modalities
  - "Put that there" paradigm

28

28

Dimension 3: Fusion

- **Higher levels: chunks from dictinct contexts**
  - Fusion of events
    - Palette
    - Drawing area

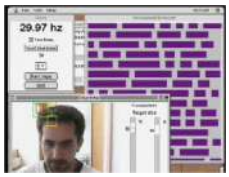


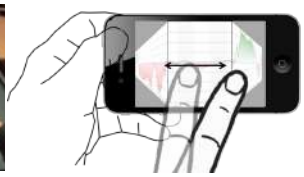
29

29

Dimension 3: Fusion

- **Higher levels: chunks from distinct contexts**

30

30

Multimodal versus multimedia

A multimodal system:

Value "Meaning" along the axis "Levels of Abstraction"

=> Four types of multimodal systems

31

31

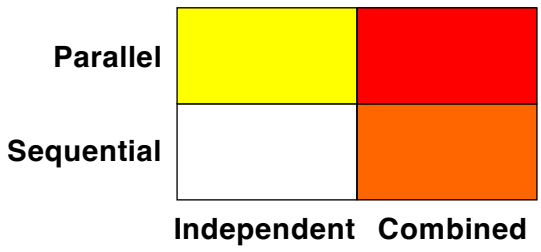
Multimodal system: four types CASE

Exclusive: (Sequential, Independent)

Alternate: (Sequential, Combined)

Concurrent: (Parallel, Independent)

Synergic: (Parallel, Combined)

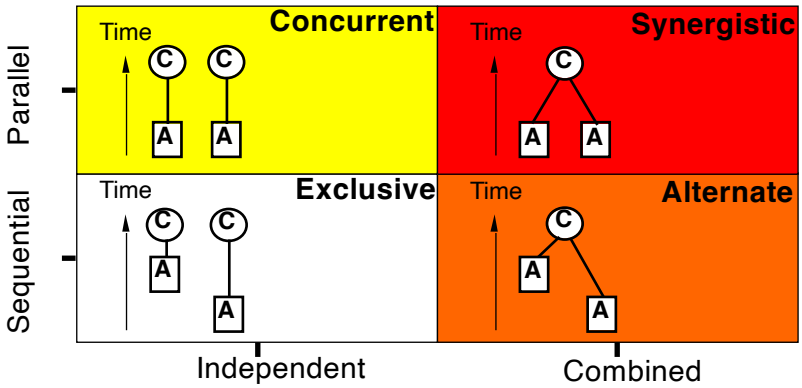


32

32

Multimodal system: four types

A multimodal system:
Value "Meaning" along the axis "Levels of Abstraction"



A User's actions
C Command, smallest fusion of user's actions that changes the system state

33

33

Multimodality: design

Underlying concepts

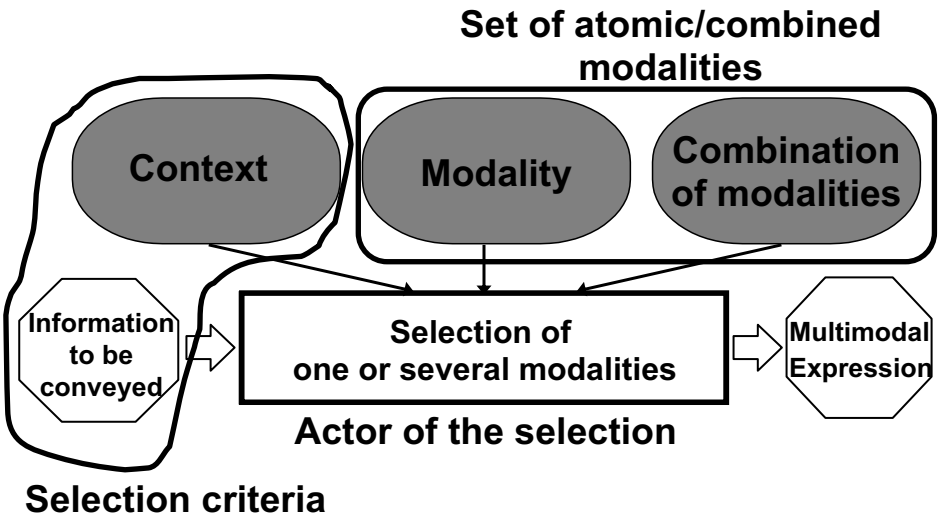
Design space

Rules of thumb, heuristics

34

34

Multimodality: Design space

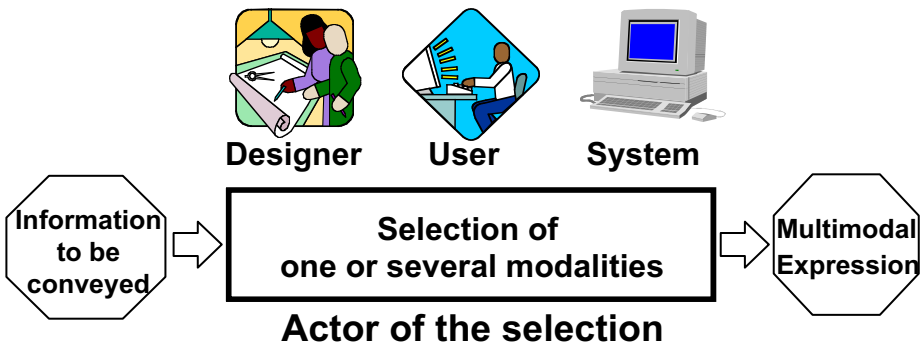


35

35

Multimodality
Actor of the selection

- Who is performing the selection



36

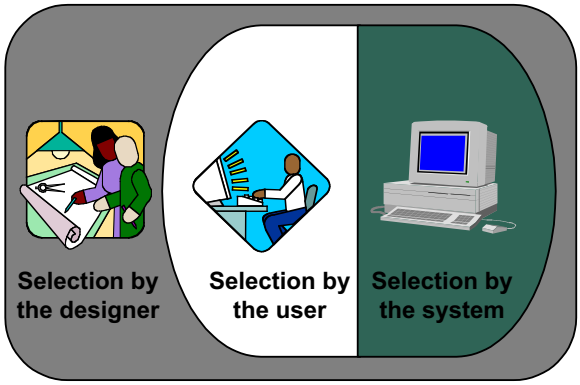
36

Multimodality
Actor of the selection

No adaptation

Adaptability

Adaptivity



37

37

Multimodality
Adaptability

Gestural modality

Speech

Embodied modality

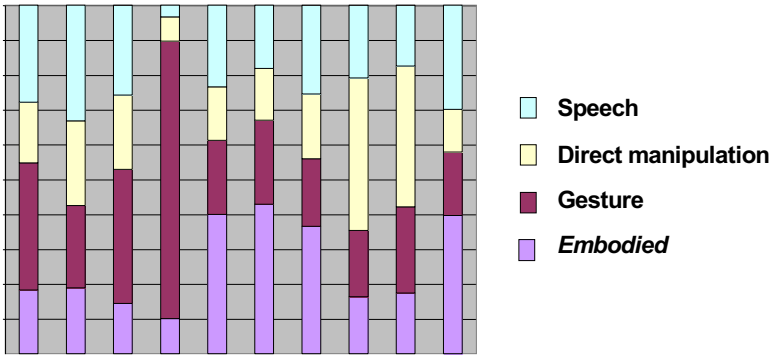
Direct manipulation

38

38

Multimodality Adaptability

- Usage of the modalities
- All sessions / All subjects



39

39

Multimodality Adaptativity

- Selection of the modalities by the system
- Context-aware systems (passive modality)

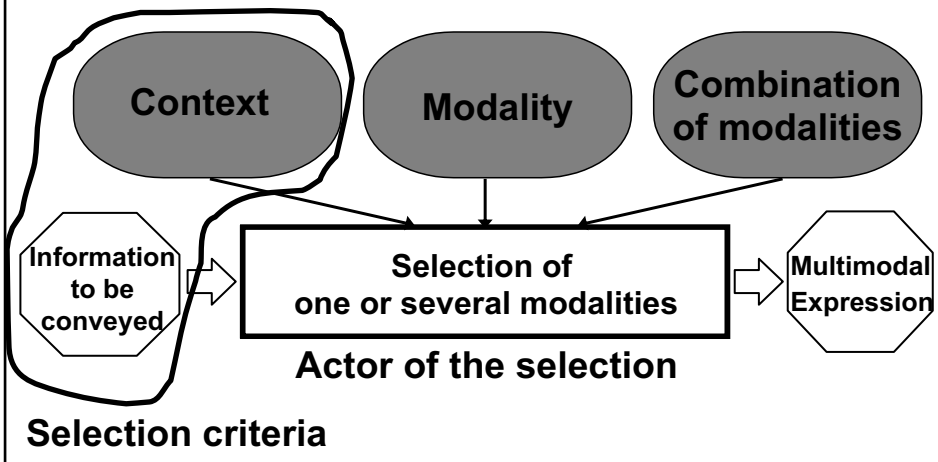
Ring

Vibration

40

40

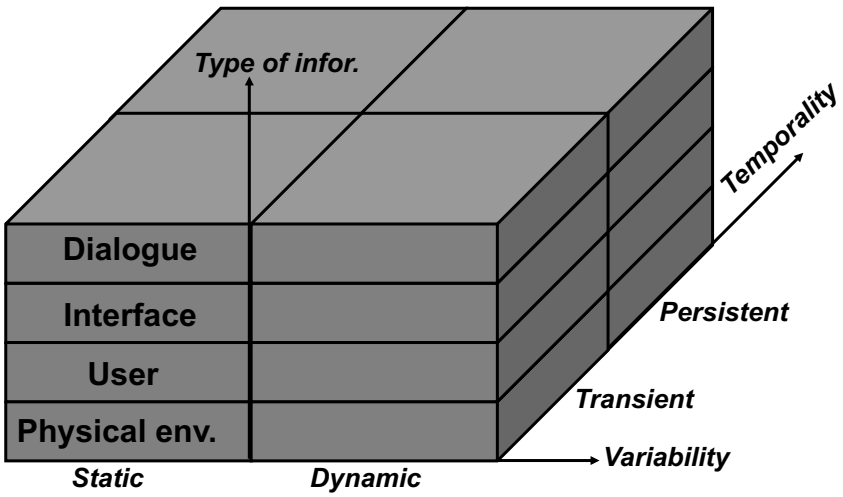
Multimodality: Design space



41

41

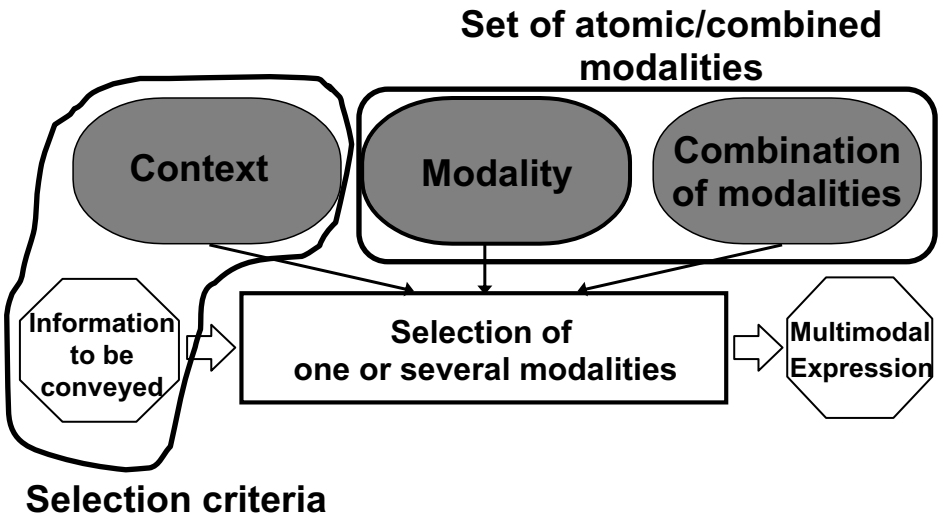
Multimodality
Selection criteria: Context



42

42

Multimodality: Design space

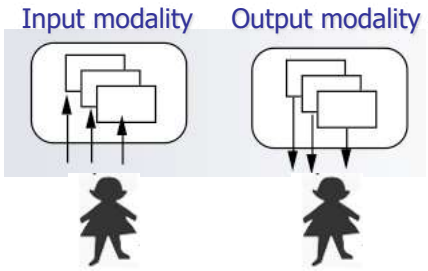


43

43

Multimodality
Characterisation of a modality

- Definition of a modality
- Modality = (device, interaction language)
  - A set of sensors (input devices) or effectors (output devices)
  - A processing facility based on a language



44

44

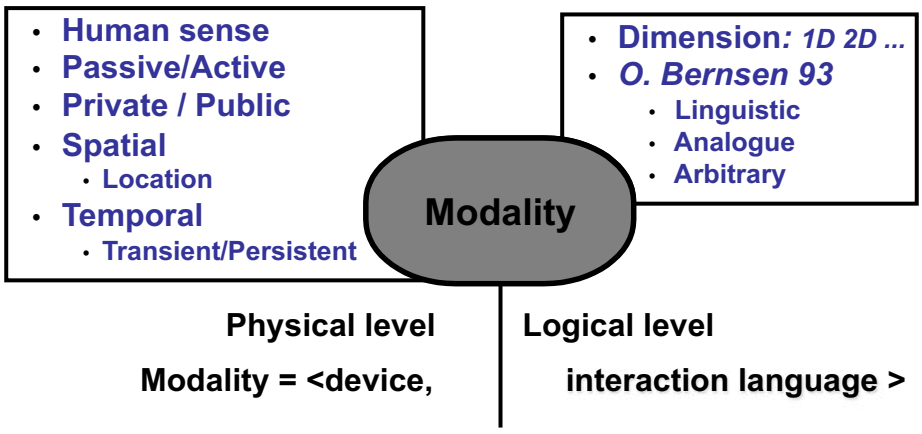
Multimodality
Characterisation of a modality

- **ACTIVE** MODALITIES
  - For inputs, active modalities are used by the user to issue a command to the computer such as a pedal to move a laparoscope in a CAS system.
- **PASSIVE - IMPLICIT** MODALITIES
  - Passive modalities are used to capture relevant information for enhancing the realization of the task, information that is not explicitly expressed by the user to the computer (PUI). For example tracking position.

45

45

Multimodality
Characterisation of a modality



46

46

Multimodality
Characterisation of a modality

47

47

Multimodality
Characterisation of a modality

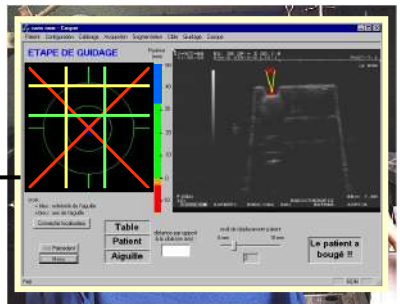
- **Physical level**
  - **Human sense: Sight**
  - **Spatial:**
    Location = operating field
  - **Temporal: Persistent**
- **Logical level**
  - **3D**
  - **Analogue**
  - **Non arbitrary**

48

48

Multimodality
Characterisation of a modality

- **Physical level**
  - **Human sense: Sight**
  - **Spatial: Location = screen**
  - **Temporal: Persistent**
- **Logical level**
  - **2D**
  - **Non Analogue**
  - **Arbitrary**



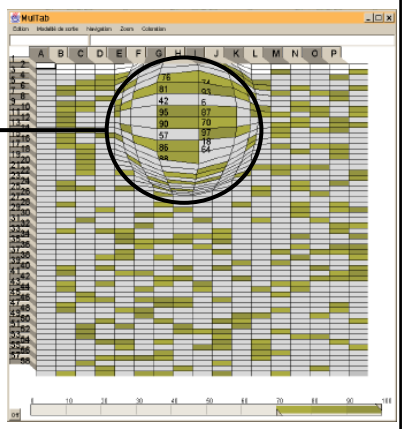
49

49

Multimodality
Characterisation of a modality

- Characterisation of a modality

- **Physical level**
  - **Human sense: Sight**
  - **Spatial: Location = screen**
  - **Temporal: Persistent**
- **Logical level**
  - **3D**
  - **Analogue**
  - **Non arbitrary**



50

50

Multimodality
Characterisation of a modality

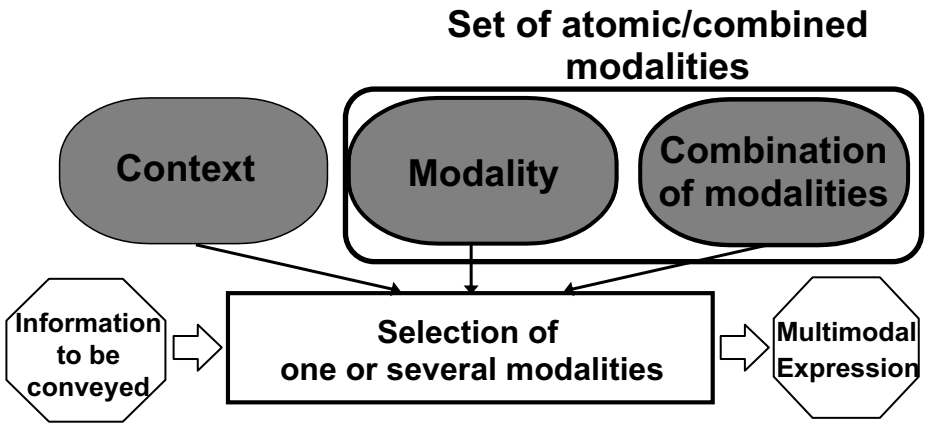
- Phycons as input modalities

- **Physical level**
  - **Human manipulation**
  - **Spatial: Location = desk**
  - **Temporal: Persistent**
- **Logical level**
  - **3D gesture**
  - **Analogue**
  - **Non Arbitrary**

51

51

Multimodality: Design space



52

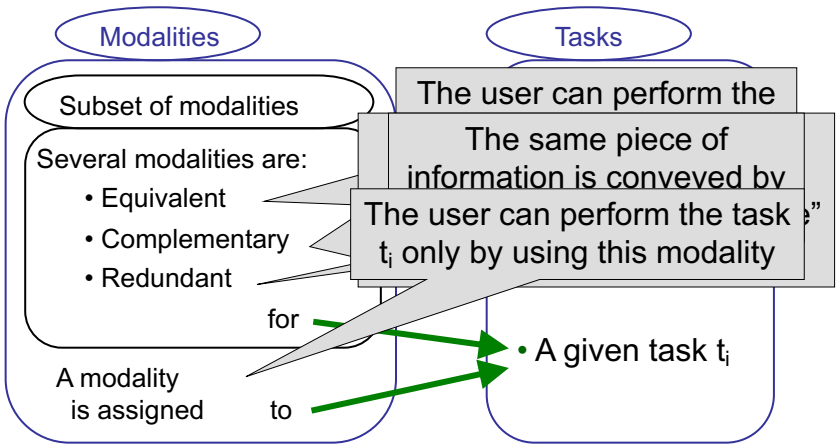
52

Multimodality
Combination of modalities

- CARE properties
  - Relationships between Devices, Interaction languages and Tasks
    - C : Complementarity
    - A : Assignment
    - R : Redundancy
    - E : Equivalence

53

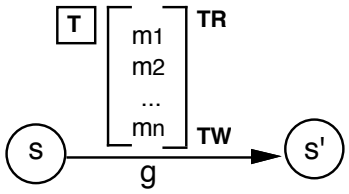
Multimodality
Combination of modalities



54

Multimodality
Combination of modalities

- The formal expression of the CARE properties relies on the notions of state, goal, modality, and temporal relationships
  - Modality mi= <d, I> = an interaction method that an agent (user, system) can use to reach a goal.
  - Goal g = a state s' that an agent intends to reach from s using modalities m_1, m_n
  - TR = temporal relationships between the use of modalities m1, .. Mn (parallelism, sequentiality, cardinality) = // | ; | 1
  - TW = temporal window within which the modalities are used
  - T = C | A | R | E = Complementarity, Assignment, Redundancy, Equivalence

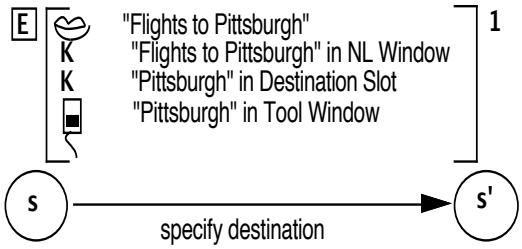


55

55

Equivalence

- Modalities of set M are equivalent for reaching s' from s, if it is necessary and sufficient to use any one of the modalities



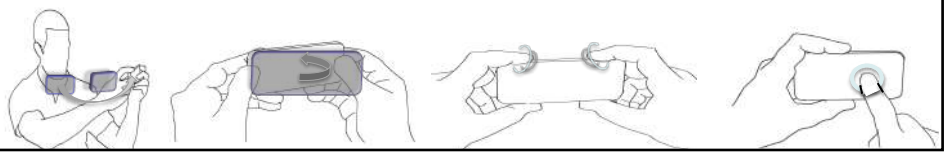
- Equivalence (s, M, s')
  ⇔
  (Card(M) >1) ∧ (∀m ∈ M Reach (s, m, s'))

56

56

Equivalence

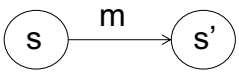
- Modalities of set M are equivalent for reaching s' from s, if it is necessary and sufficient to use any one of the modalities



57

Assignment

- In contrast to equivalence, assignment expresses the absence of choice.
- Modality m is assigned in state s to reach s', if no other modality can be used to reach s' from s



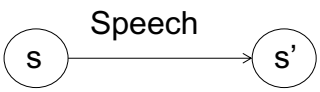
- Assignment (s, m, s')
  $\Leftrightarrow$
  Reach (s, m, s')
  $\wedge$ ($\forall$ m'$\in$ M. Reach(s, m',s') $\Rightarrow$ m'=m)

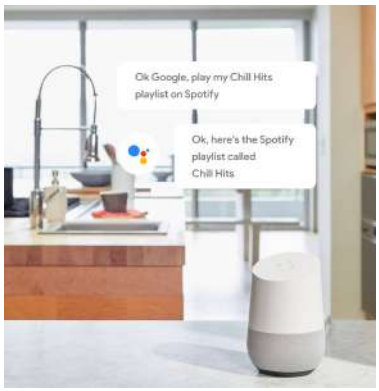
58

58

Assignment

- Modality m is assigned in state s to reach s', if no other modality can be used to reach s' from s



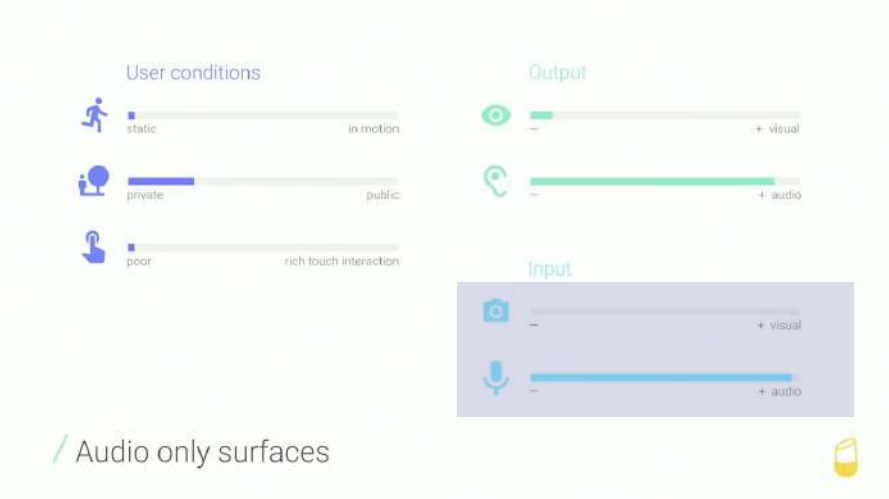


https://store.google.com/uk/product/google_home

59

59

Assignment

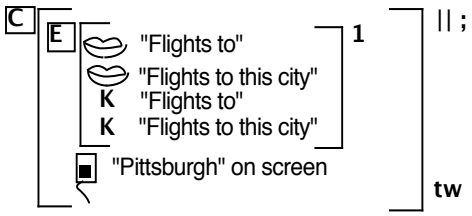


60

60

Complementarity

- Modalities of a set M must be used in a complementary way to reach state s' from state s within a temporal window, if all of them must be used to reach s' from s, i.e., none of them taken individually can cover the target state. (fusion required)



61

61

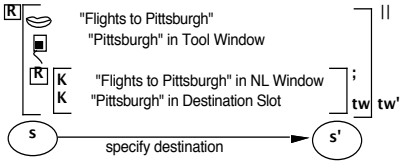
Complementarity

62

62

Redundancy

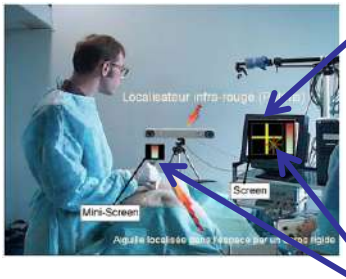
- Modalities of a set M are used redundantly to reach state s' from state s, if they have the same expressive power (they are equivalent) and if all of them are used within the same temporal window, tw.

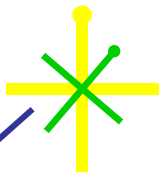


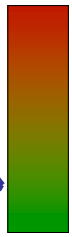
Redundancy

Mark

https://www.cybertek.fr/peripherique-de-jeu/thrustmaster/33079.aspx







Multimodality
Combination of modalities

- CARE : relationships between modalities and tasks

=> Semantic relationships

- Adding the temporal aspect when two modalities are used

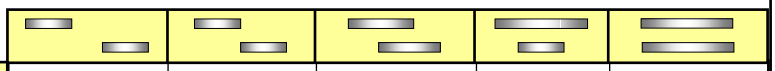
- **2 aspects: temporal and semantic**
- **5 schemas: [Allen 83]**

65

65

Multimodality:
Combination of modalities

Combination schemas

Combination aspects					
Temporal	Anachronism	Sequence	Concomitance	Coincidence	Parallelism
Semantic	Concurrency	Complementarity	Complementarity & Redundancy	Partial Redundancy	Total Redundancy

66

66

Multimodality: Combination of modalities

Temporal	Anachronism	Sequence	Concomitance	Coincidence	Parallelism
Semantic	Concurrency	Complementarity	Complementarity & Redundancy	Partial Redundancy	Total Redundancy

	Independent	Combined
Parallel	**Concurrent**	**Synergic**
Sequential	**Exclusive**	**Alternate**

67

67

Multimodality: Combination of modalities

Temporal	Anachronism	Sequence	Concomitance	Coincidence	Parallelism
Semantic	Concurrency	Complementarity	Complementarity & Redundancy	Partial Redundancy	Total Redundancy

	Independent	Combined
Parallel	**Concurrent**	**Synergic**
Sequential	**Exclusive**	**Alternate**

68

68

Multimodality: Combination of modalities

Temporal	Anachronism	Sequence	Concomitance	Coincidence	Parallelism
Semantic	Concurrency	Complementarity	Complementarity & Redundancy	Partial Redundancy	Total Redundancy

	Independent	Combined
Parallel	Concurrent	Synergic
Sequential	Exclusive	Alternate

69

69

Multimodality: Combination of modalities

Temporal	Anachronism	Sequence	Concomitance	Coincidence	Parallelism
Semantic	Concurrency	Complementarity	Complementarity & Redundancy	Partial Redundancy	Total Redundancy

	Independent	Combined
Parallel	Concurrent	Synergic
Sequential	Exclusive	Alternate

70

70

Multimodality: Combination of modalities

Adding the **spatial aspect** when two modalities are used

Combination schemas

Combination aspects					
Temporal	Anachronism	Sequence	Concomitance	Coincidence	Parallelism
Spatial	Separation	Adjacency	Intersection	Overlaid	Collocation
Semantic	Concurrency	Complementarity	Complementarity & Redundancy	Partial Redundancy	Total Redundancy

71

71

Multimodality: Combination of modalities

Combination schemas

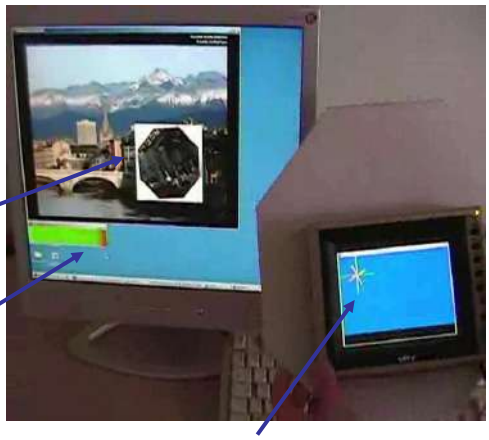
Combination aspects					
Temporal	Anachronism	Sequence	Concomitance	Coincidence	Parallelism
Spatial	Separation	Adjacency	Intersection	Overlaid	Collocation
Semantic	Concurrency	Complementarity	Complementarity & Redundancy	Partial Redundancy	Total Redundancy

72

72

Multimodality: Combination of modalities

- Puzzle



M1 = <screen, 2D image>

M2 = <screen, color>

M3 = <mini-screen, crosses>

73

73

74

74

Multimodality: Combination of modalities

- Puzzle

75

75

Multimodality: Combination of modalities

- Combination of
  M2 = <screen, color> and
  M3 = <mini-screen, crosses>

Combination schemas

Combination aspects					
Temporal	Anachronism	Sequence	Concomitance	Coincidence	Parallelism
Spatial	Separation	Adjacency	Intersection	Overlaid	Collocation
Semantic	Concurrency	Complementarity	Complementarity & Redundancy	Partial Redundancy	Total Redundancy

76

76

Multimodality
Combination of modalities

- CARE properties
  - C : Complementarity
  - A : Assignment
  - R : Redundancy
  - E : Equivalence

Combination schemas

Combination aspects

Temporal	Anachronism	Sequence	Concomitance	Coincidence	Parallelism
Spatial	Separation	Adjacency	Intersection	Overlaid	Collocation
Semantic	Concurrency	Complementarity	Complementarity & Redundancy	Partial Redundancy	Total Redundancy

77

77

Multimodality
Combination of modalities

Task: Update position & orientation of User

Tasks: Manipulation of Note (create, pick, and remove)

CARE combination

Complementarity 2

Complementarity 1

Redundancy / Equivalence

Languages

3D orientation (radians)

3d location

Mouse Notes commands

Speech Notes commands

Devices

Magnetometer

Localization sensor

Mouse

Microphone

78

78

Multimodality
Combination of modalities

79

79

Multimodality
Combination of modalities

- CARE properties
- TYCOON design space

80

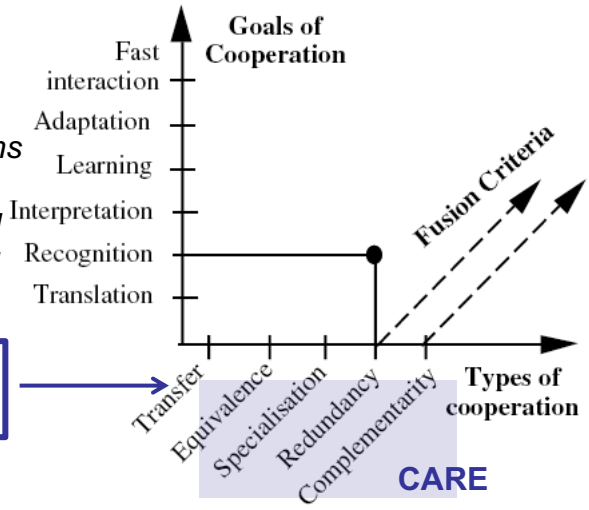
80

Multimodality
Combination of modalities

- TYCOON

When several modalities cooperate by transfer, this means that a chunk of information produced by a modality is used by another modality

TRANSFER

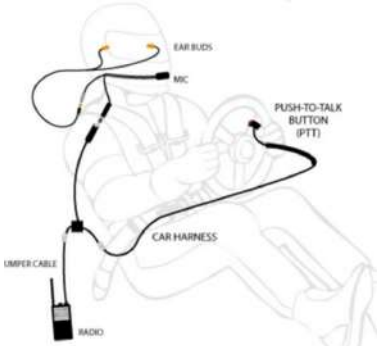


81

81

Multimodality
Transfer of modalities

- One modality is used to activate another modality
- Push-To-Talk
- <button, press/release> *(active modality)*
  - <microphone, speech command> *(active modality)*



https://www.sampsonracing.com/2-Way-Hole-Mount-PTT-with-Mounting-Bracket-p/p0520.htm

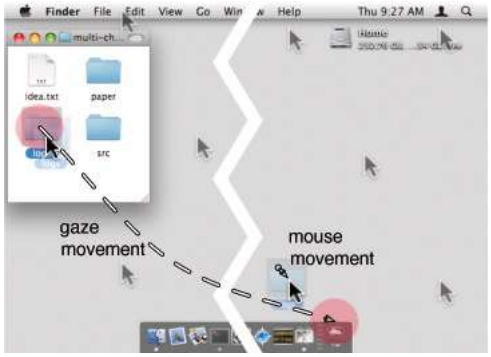
82

82

Multimodality
Transfer of modalities

- Eye-tracking is used to select the cursor manipulated by the mouse
- <eye-tracking, 2D position> *(passive modality)*

  ➡ <mouse, direct manipulation> *(active modality)*



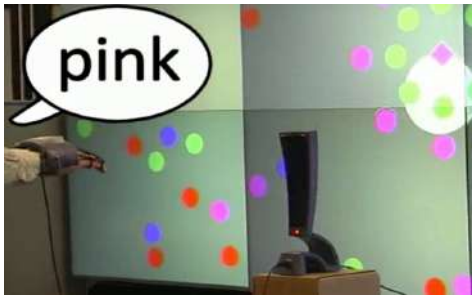
83

83

Multimodality
Transfer of modalities

- Speech is used to modify the behavior of the pointing technique
- <Microphone, name of a color> *(active modality)*

  ➡ <3D gesture, pointing-bubble cursor> *(active modality)*



84

84

Multimodality: design

Underlying concepts

Design space

Rules of thumb, heuristics

85

85

Multimodal input/output

- *Supporting documents: heuristics.pdf*
- A set of multimodal design principles that are founded in perception and cognition science

- Four general areas
  - Designing multimodal input and output
  - Adaptivity
  - Consistency
  - Feedback
  - Error prevention/handling

86

86

Multimodal input/output

- *Supporting documents: heuristics.pdf*
- Designing multimodal input and output
  - Match output to acceptable user input style
    - if the user is constrained by a set grammar, do not design a virtual agent to use unconstrained natural language
- Adaptivity
  - Multimodal interfaces should adapt to the needs and abilities of different users, as well as different contexts of use. Dynamic adaptivity enables the interface to degrade gracefully by leveraging complementary and supplementary modalities according to changes in task and context.
    - Allowing gestures to augment or replace speech input in noisy environments, or for users with speech impairments

87

Multimodal input/output

- *Supporting documents: heuristics.pdf*
- Consistency
  - System output independent of varying input modalities
    - the same keyword provides identical results whether user searches by typing or speaking
- Feedback
  - Users should know which modalities are available to them
- Error Prevention/Handling
  - If an error occurs, permit users to switch to a different modality

88

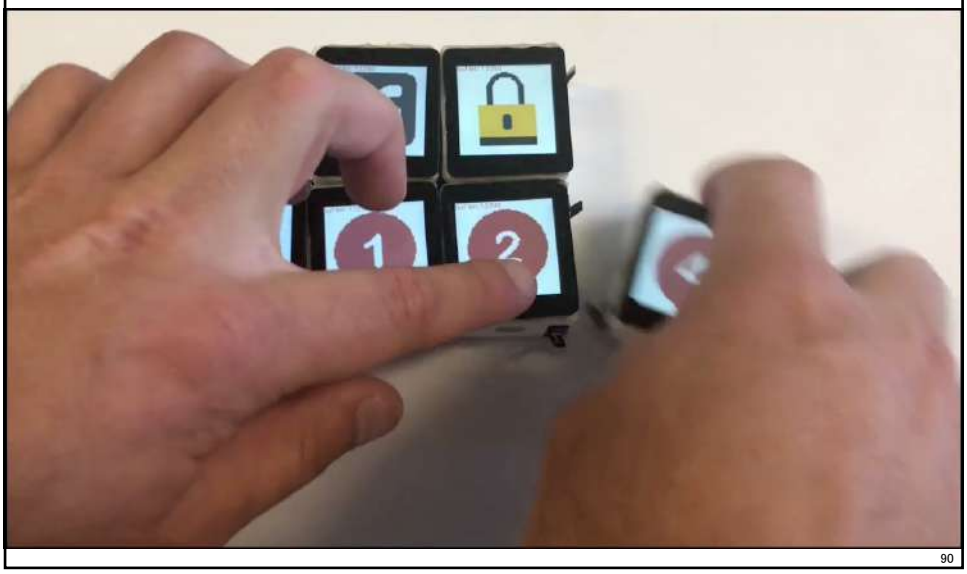
Design: Main points

- Design space for multimodal interaction
  - Characteristics of a modality
  - Composition space
- Mapping of functionalities onto modalities not always straightforward
  - Few guidelines
  - Experimental study

89

89

One research avenue: reconfigurable device

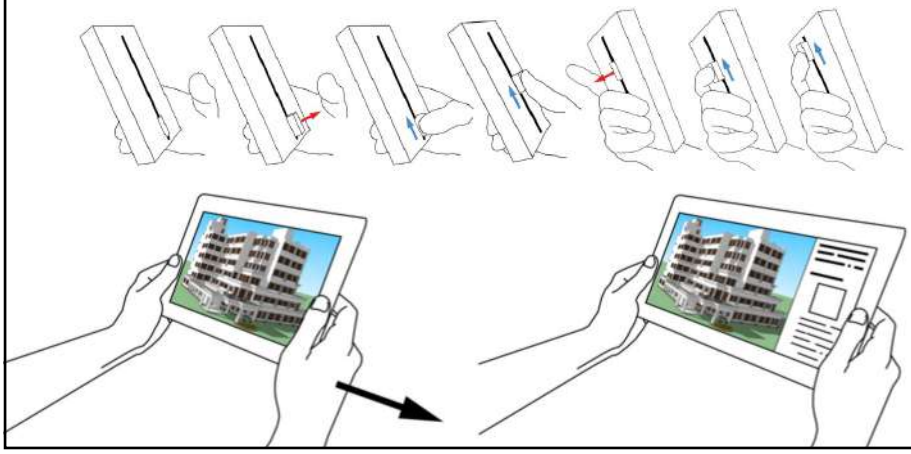


90

90

One research avenue: shape-changing UI

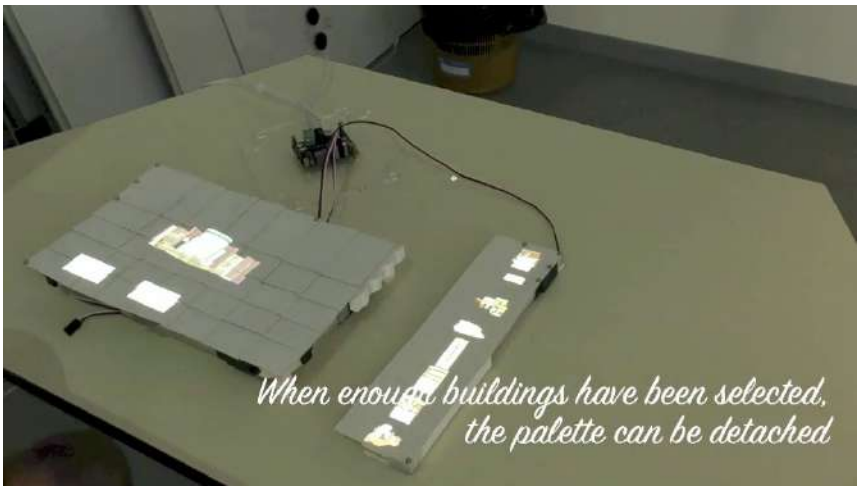
- Modality <device, language> with a deformable device



91

91

One research avenue: shape-changing UI



92

92

Readings

- Bernsen, N. Modality Theory in support of multimodal interface design. Proceedings of Intelligent Multi-Media Multi-Modal Systems, (1994), pp. 37-44
- Bouchet, J., Nigay, L., Ganille, T. ICARE Software Components for Rapidly Developing Multimodal Interfaces. Proceedings of ICMI'04, ACM Press, pp. 251-258
  http://iihm.imag.fr/en/publication/
- Coutaz, J., et al. Four easy pieces for assessing the usability of multimodal interaction: The CARE properties, Proceedings of Interact'95, Chapman&Hall, pp. 115-120
  http://iihm.imag.fr/en/publication/
- Martin, J. C. TYCOON: Theoretical Framework and Software Tools for Multimodal Interfaces. Intelligence and Multimodality in Multimedia Interfaces, AAAI Press (1997)
- Nigay, L., Coutaz, J. The CARE Properties and Their Impact on Software Design. Intelligence and Multimodality in Multimedia Interfaces, (1997)
  http://iihm.imag.fr/en/publication/
- Vernier, F,. Nigay, L. A Framework for the Combination and Characterization of Output Modalities, Proceedings of DSV-IS2000, Springer-Verlag, pp. 32-48
  http://iihm.imag.fr/en/publication/

93