# Localization of media-rich interactive ads

Mario De Bortoli (London)
Fabio Minazzi (Milan)

## *Abstract*

At a time when media convergence has become a reality online advertising agencies are more and more often confronted with problems that used to affect the TV-based advertising agencies. With the spreading of streaming video, podcasts and interactive video technologies, all of a sudden websites and online ads are full of life, but this requires the rethinking of the text-centric localization processes, to move the focus to the audio/visual elements, that introduce an additional parameter to be considered during the localization process: media synchronization in time.

The typical fordistic localization process where the text is translated first and then it is fed in the various media is not suitable for this type of products. Multidimensional translation has to be combined with interactive design and media localization techniques, to provide effective multilingual communication, able to stand the quality of traditional audiovisual products.

And it is precisely the issue of changing localization processes within the media-rich and online advertising industry that the authors try to tackle by sharing their day to day experiences and acquired knowledge to work out a model for designing the most appropriate process for the localization of international campaigns and online content in the age of media convergence.

The paper gives a brief overview of how the online communication is localized, illustrating the most common processes used in the industry. The authors move on to explain how the media-rich contents is changing the face of localization in terms of resources, skills and processes needed to be integrated in the workflow, to address the ever-increasing use of interaction, subtitling, video, dubbing and voiceovers.

A theoretical framework is proposed, where media-rich scenes are divided into 5 main components (audio, video, graphics, static and dynamic on screen text , interactive items), which are then weighted by relating them with 4 main types of constraint (space, time, cultural functional).

Once 3 additional parameters (time, cost, and amounts) are introduced, it is possible to define the driving localization components, that are those components that shape the localization process. Examples of the application of this model are provided.

The aim of this work is to provide a useful theoretical framework to communication designers, localization managers, local marketing managers, and more generally to the people involved in the production and localization of global communication products.

## *Introduction*

In a global communication world, industrialization is a must. An appropriate understanding of the production and localization processes is the key to give the consumer an exciting, consistent and interactive experience, while containing costs for the advertiser.

The paper gives a brief overview of how the online communication is localised, illustrating the most common processes used in the industry. The authors move on to explain how the media-rich contents is changing the face of localization in terms of resources, skills and processes needed to be integrated in the workflow, to address the ever-increasing use of interaction, subtitling, video, dubbing and voiceovers.

A theoretical framework is proposed, where media-rich scenes are divided into 5 main components (audio, video, graphics, static and dynamic on screen text , interactive items), which are then weighted by relating them with 4 main types of constraint (space, time, cultural functional). Once 3 additional parameters (time, cost, and amounts) are introduced, it is possible to define the driving localization components, that are those components that shape the localization process.

## Multimedia scenes: a combination of 5 text-containing components and 4 constraints

In principle all multimedia scenes are built with 5 multimedia components which can contain text:
1. Audio: the spoken text
2. Video: subtitles and other overlaid graphic text
3. Raster graphics: the static text contained in the pictures and screen shots
4. Software - on screen text: the text dynamically displayed by the content presentation system
5. Software - interactive elements: the text contained into the elements that are devoted to interacting with the multimedia product, i.e. buttons, menus, dialogues, input fields.

Except for the audio component, all the others form the visual scene on screen.

The localization of each component requires a separate *production line*, and specific *quality controls* (QA) defined according to the type of constraints affecting each component.

When localizing each component, the text must be translated taking into account **4 types of constraints**

1. Space constraints
2. Time constraints
3. Cultural constraints
4. Functional constraints

The importance of each constraint varies from product to product, according to the combination of components in each scene, and causes the role (weight) of each component to change in the design of the production method.
The table below outlines the influence of each constraint on the translation of the text contained in each component.

**1) Space constraints**

| Space constraint influence on translation of: | | | | |
|---|---|---|---|---|
| Audio | Video | Graphics | On screen text (software) | Interface elements (software) |
| Usually not influenced by this constraint. | Subtitles: very space-constrained (typical: 36 chars on max 3 lines). Graphics: very limited. May need insertion between source lines. This is often a critical component. | Static visual elements always have a limited amount of space for displaying. | Dynamic text may be displayed in boxes that allow for expansion or scroll, this influences the strength of the constraint. | Short sets of specific words can become much longer or shorter when translated, and consistency is a must. This is often a critical component. |

**2) Time constraints**

| Time constraints influence on translation of: | | | | |
|---|---|---|---|---|
| Audio | Video | Graphics | On screen text (software) | Interface elements (software) |
| Sound is usually synchronized with the action on screen: translation should account for total length of presentation and intermediate cue points as well | Amount of text displayed depends on the amount of time available for displaying it | Influenced if the audio or video text refers explicitly to the content of the graphics | Influenced if the audio or video text refers explicitly to the content of the on screen text | These are usually not influenced |

**3) Cultural constraints**

| Cultural constraints influence on translation of: | | | | |
|---|---|---|---|---|
| Audio | Video | Graphics Influenced in terms of: | On screen text (software) Influenced in terms of: | Interface elements (software) Influenced if |
| Influenced in terms of: <br> - speaker's pace <br> - actor's gender <br> - text (may need rewriting) <br> - accents | Subtitles are influenced in terms of: <br> - subtitling style <br> - text (may need rewriting) | - content layout <br> - screen shoots (for products that contain references to screen action) <br> - fonts <br> - text (may need rewriting) | - content layout (may need resize and re-position) <br> - text (may need rewriting) | software functions vary from locale to locale |

**4) Functional constraints**

| Functional constraint influence on translation of: | | | | |
|---|---|---|---|---|
| Audio text | Video text | Graphics text | On screen text (software) | Interface elements (software) |
| Influenced only if the text references functional elements | The text appearing in a video is very limited (typical values 36 chars on max 3 lines) <br><br> In case the video text coincides with part of the on screen text, the on screen text length is affected by the video text | Influenced if the on screen text refers explicitly to the content of the graphics | Influenced if the text refers to the functions of the software | Influenced in terms of text that needs rewriting to account for the localized functions of the software |
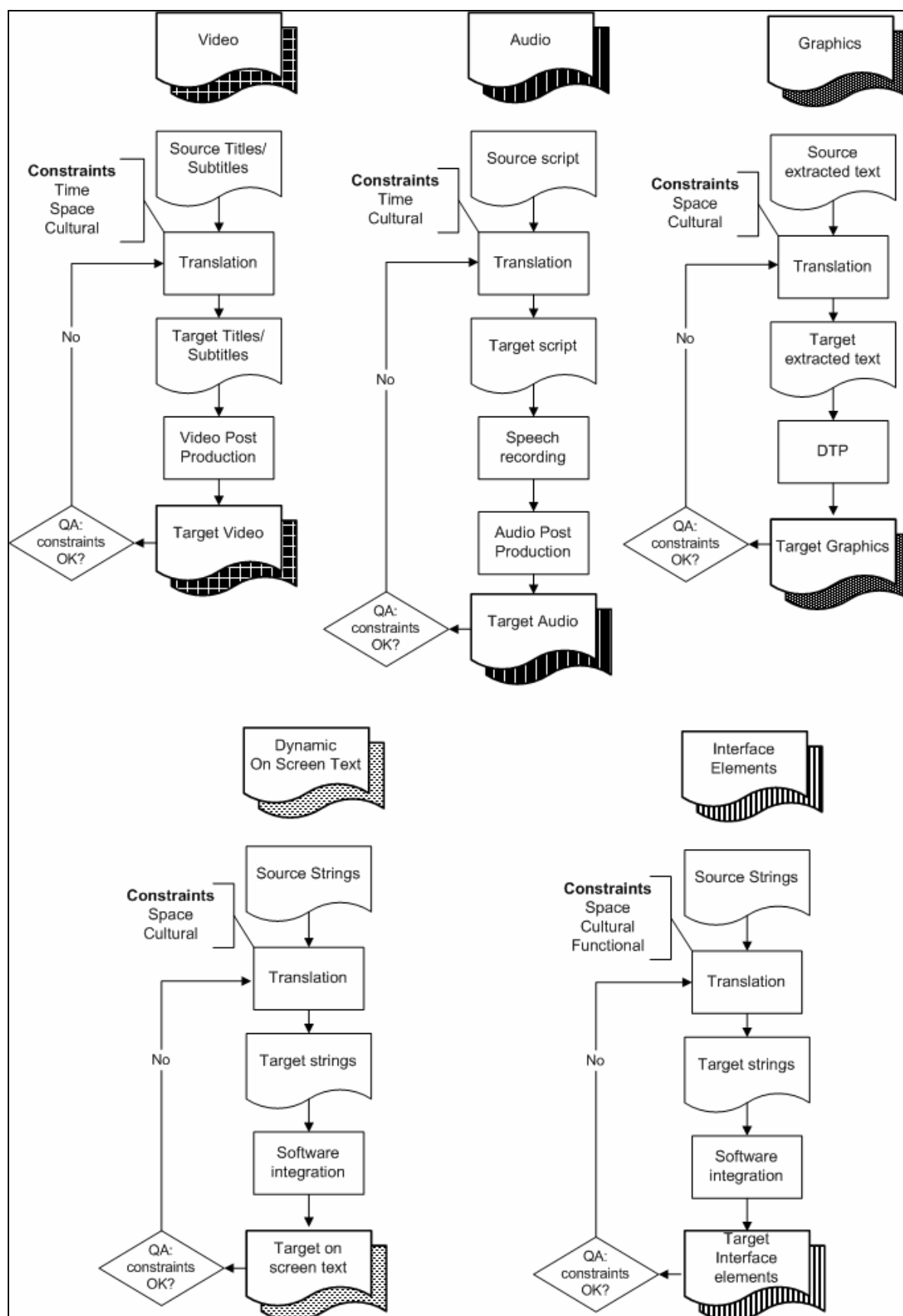
Figure 1: components, production lines and constraints

Figure 1 describes in a simplified way, the production lines for localizing each component.
The influence of the constraints on each component is also displayed.

## *Component-specific vs. cross-component constraints*

Let's now look at the influence of each component on the others.
We call *cross-component constraints* the constraints that apply on one component because of another component.

Figure 2 shows the typical result of a localization project carried out without taking into account the cross-component constraints.
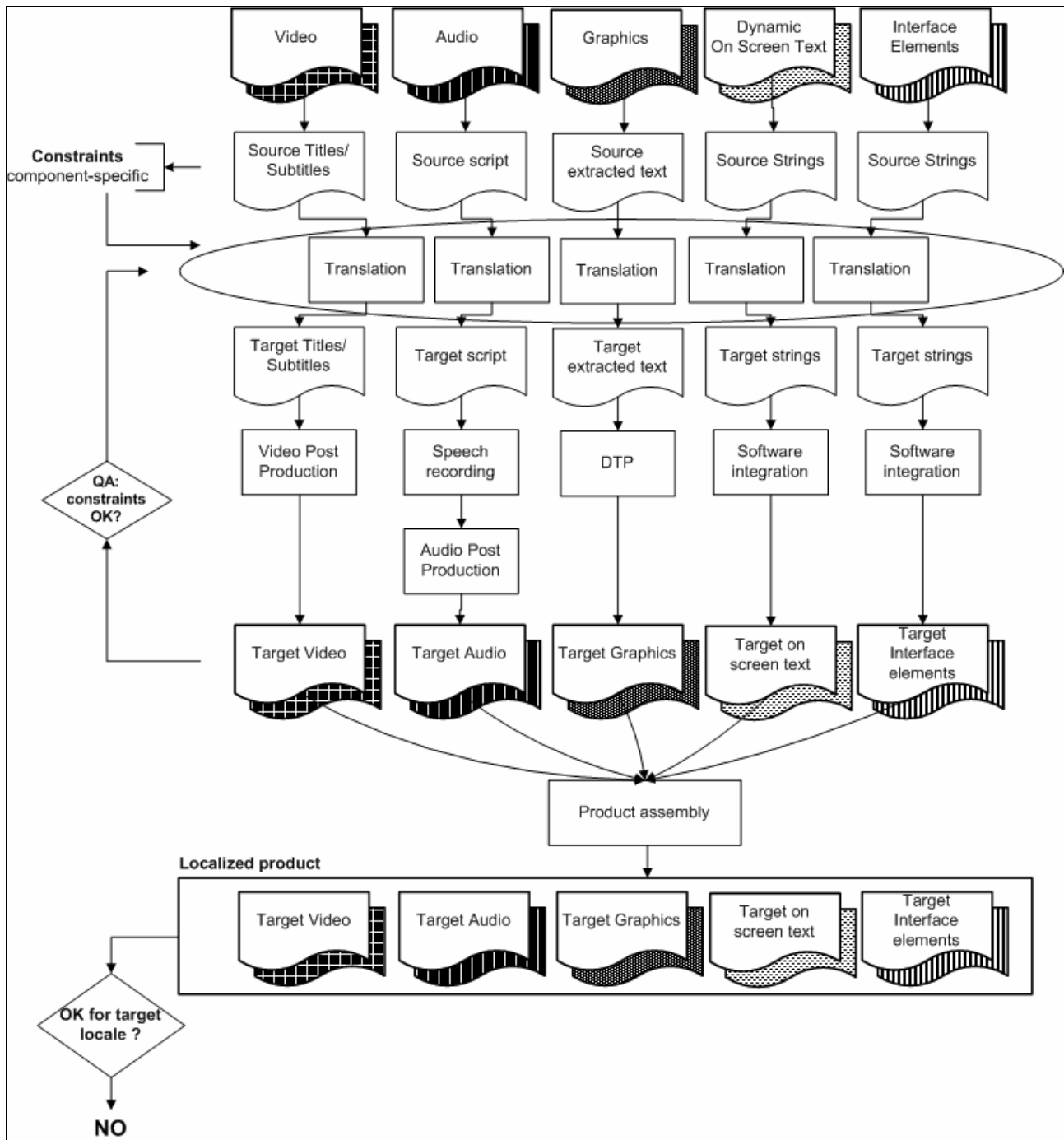


Figure 2. Lack of *cross-component constraint* analysis results in inadequate localization

An example can be useful to better explain the concept of cross-component constraint.
In a traditional film the time constraint on the video influences directly the translation of the audio script and subtitles, and that situation is the same for all films.

**Film**
*"I wish you were here to see Jane's wedding, she was absolutely wonderful"*

Time constraint of 2.5 seconds on screen, forces to change the translation for the subtitle.
Translated
*"I wish you were here for the wedding. Jane was wonderful"*

**Multimedia ad**
*"If you select the new Verify Space for My Account feature scrolling the menu, you check the available space for your messages."*

while translating it is not possible to shorten this sentence without causing a loss of part of the original message. Each bolded element has in fact a specific function related to the way the presentation  or the object described work.

*"If you **select** the **new Verify Space for My Account** feature **scrolling** the **menu**, you check the **available space** for your **messages**."*

As often happens in a multimedia products, the sentence refers to elements on the screen, and the message is an instruction built without redundance.

Multimedia products, by their very nature, feature sets of constraints that change from product to product. For example in one product the space constraints on the graphics can turn into a constraint on the audio script, while in another one the translation of the interactive elements affects the translation of the subtitles.
This entails that either during the design of the product or campaign, or at the very beginning of the localization, a detailed analysis has to be performed in order to identify the constraints and properly plan the localization process.

Figure 3 shows a case where a good cross-component analysis leads to a successful localization project. The analysis shows that the audio script and the interactive elements influence the translation of other components. The relevant constraints applied during the translation stage, and the result obtained after the product assembly, become suitable for the target locales.
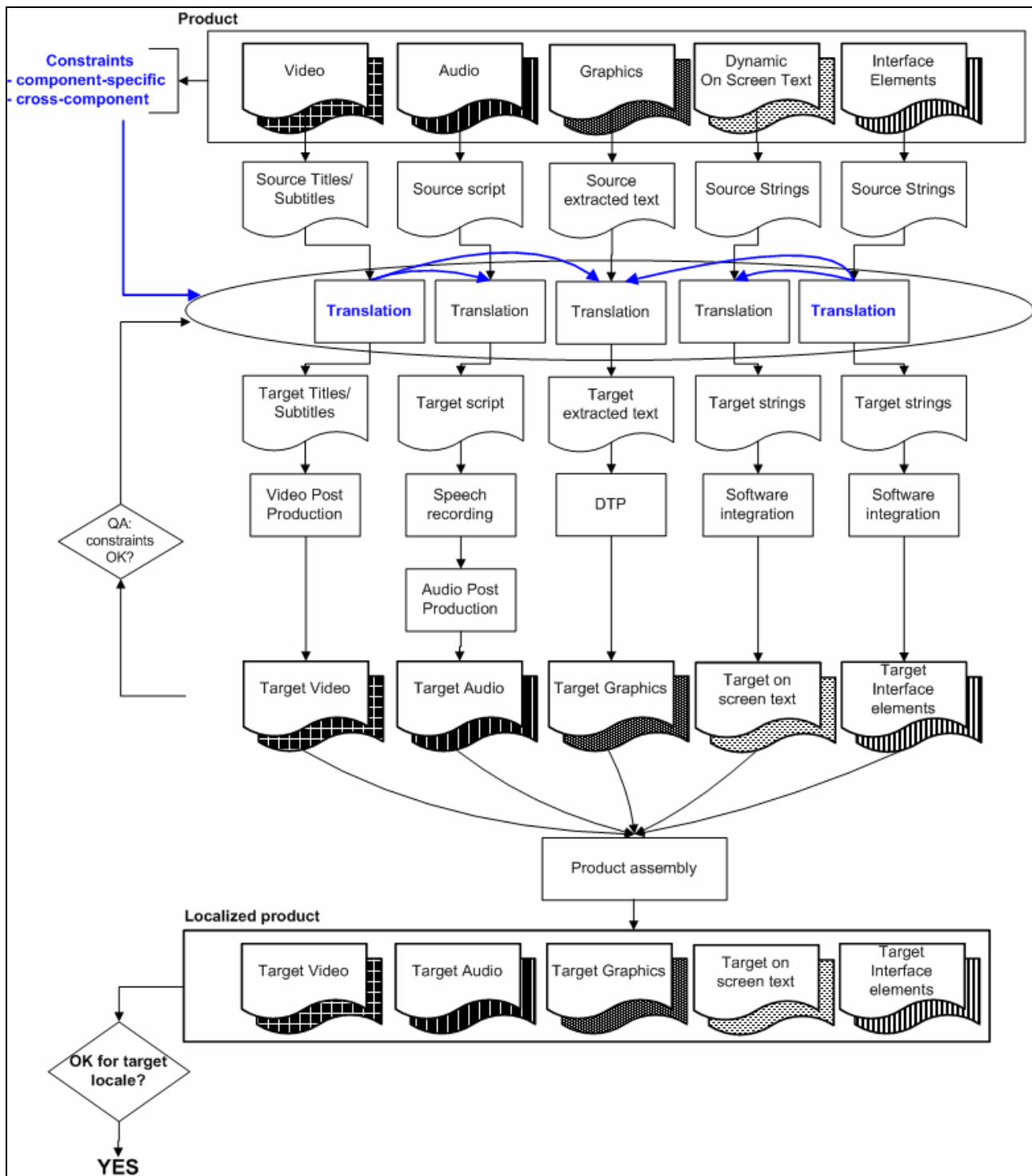
Figure 3. Good *cross-component constraint* analysis results in successful localization

It is to be noticed that these constraints trigger a set of QA activities that are to be performed on the assembled product.

It is to be noticed that the terms multimedia and media-rich contents both used in this paper, refer to the same concept: not a simple assembly of media that are provided to the user, rather, a combination of contents with complex cross-references and functional relationships.

## *The role of costs, time, amount*

In order to complete the picture we need to account for three additional parameters: **cost, time, amounts**. Each one of these parameters strongly influences the localization process:

1. *cost of each working step* in the production line

Some steps much more expensive than others by their very nature, e.g. audio recording, video editing, software engineering. As an example: once the editing of the video is completed, the cost of adding to a video a single string that has been changed/overlooked is much higher than the cost of simply re-translating it.

2. *time* involved in each working step

Some steps are *slow to activate*. For example the booking of actors for a recording of a single string can take days if the actor is on vacation. Some steps are *slow to be performed* as a result of lengthy manual operations like the manual insertion of strings in a set of layered pictures.

3. *amount* of content to be processed

The amount is always an important parameter (this has an influence on both cost and time). For example if the translated text is placed in thousands of pictures, the DTP step can become the bottleneck of the entire localization process.

## *Defining priorities in the localization process: the driving localization components*

As a result of the constraints analysis and cost/time factors we can now define the **driving localization components**: those components that due to the influence of their constraints, and of the related cost/time/amounts factors, shape the entire localization process.

Traditionally when defining the localization process usually the order of translation is done first and then, with the translated text all the media are populated.
This model does not work, in many multimedia localization cases where the order for translation is often driven by the most time-consuming task.

In this example the product to be localized features some interactive elements that generate constraints to the on screen dynamic text and the audio text as well.
We analyse two localization plannings that look equivalent, but indeed are not.
In the first localization process (see figure 4) before starting and analysis of the contents showed the cross-component constraints.
The interactive elements are outlined to be the *driving localization component* because of the constraints that are generated on other components.
Another driving localization component is the audio. In fact, the process for recording the audio is estimated to be longer than the time for integrating the on screen dynamic text so it must start first, so that  the other translations can proceed while the audio is being produced.
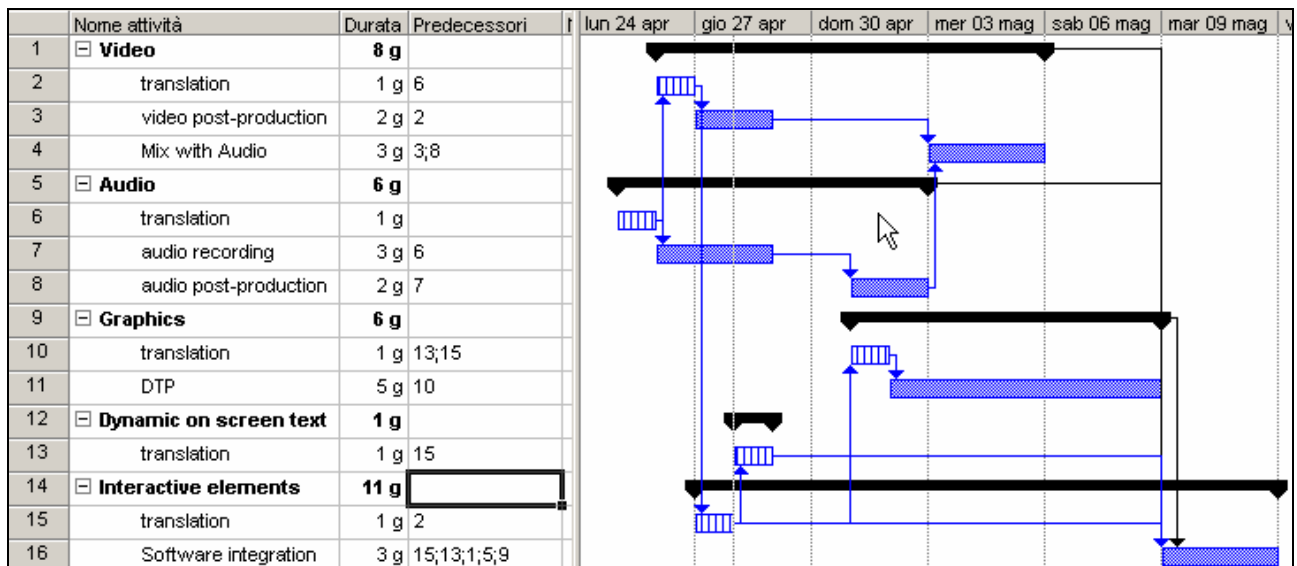The total project duration following the localization process 1 is 11 days.

Figure 4. schedule for localization process 1

If a different localization process is applied (le't s call it process 2), the most time consuming activities (audio and video) are started first, which makes sense in order to minimize the production time. Apparently the project duration "is the same" (see figure 5).
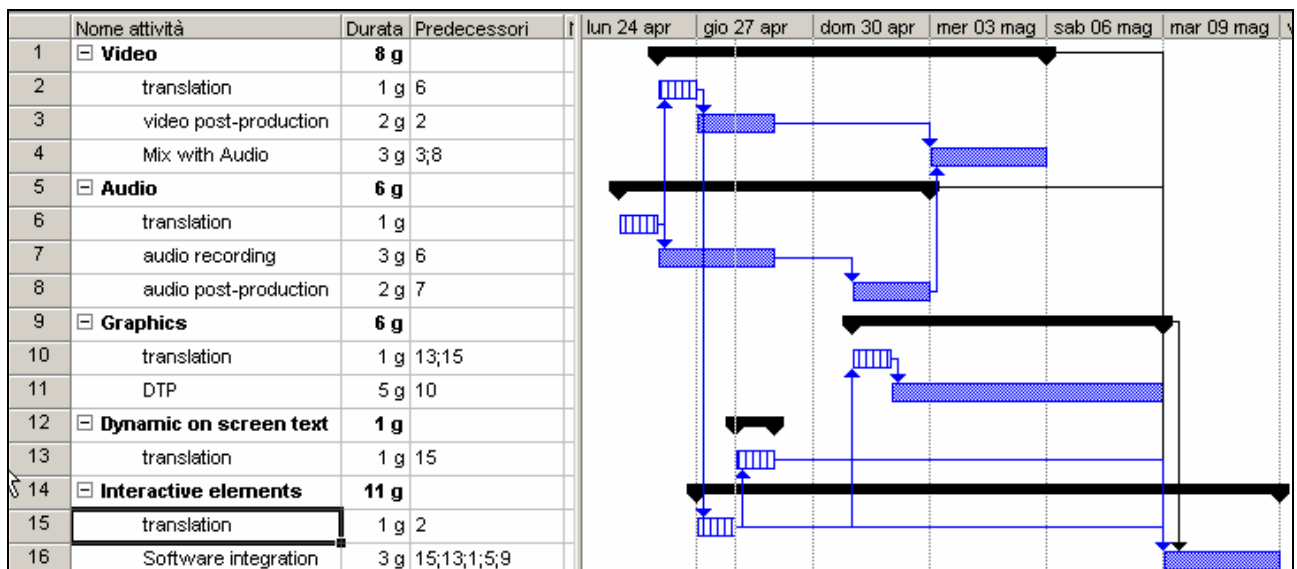


Figure 5. schedule for localization process 2 (theoretical)

But the cross-component constraints have not been taken into account.
Once the audio is integrated with the video and the remaining components in the product, the cross-component constraints become apparent, and this causes a redo of the audio translation and recording  after the integration.
The result is that if one selects process 2 in place of process 1, the real project schedule ends up to be 18 days,  50% longer than process 1 (see figure 6), and in additional costs. This effect is especially dangerous when several languages are to be localized in parallel.

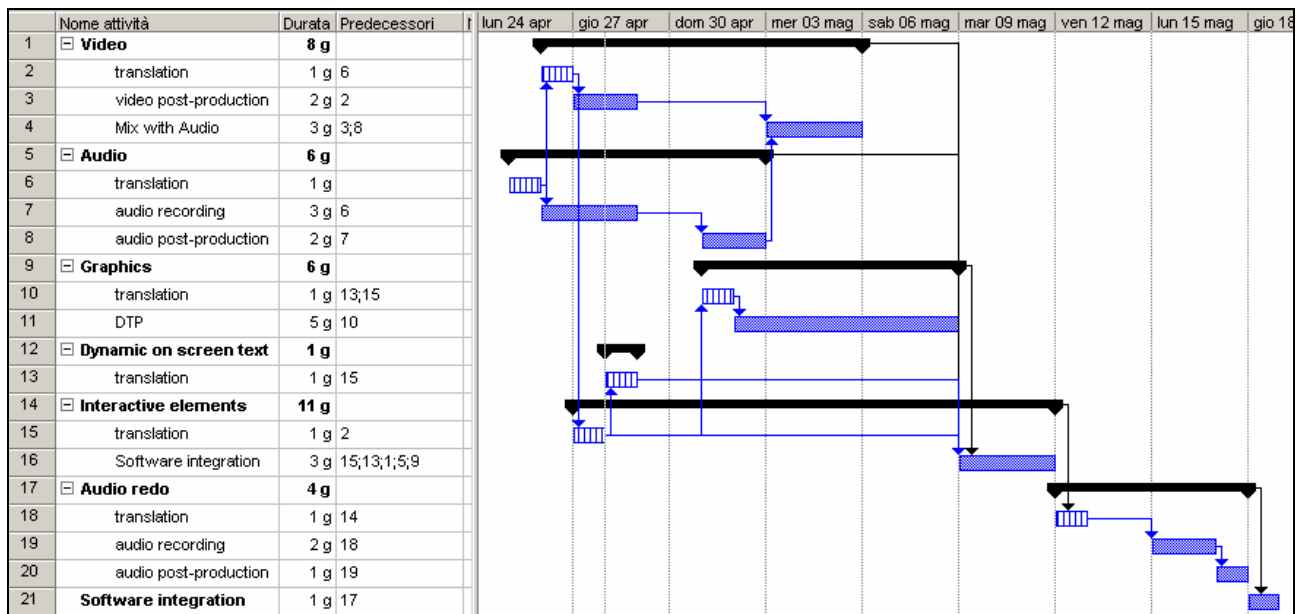| | Nome attività | Durata | Predecessori |
|---|---|---|---|
| 1 | ⊟ **Video** | **8 g** | |
| 2 | translation | 1 g | 6 |
| 3 | video post-production | 2 g | 2 |
| 4 | Mix with Audio | 3 g | 3;8 |
| 5 | ⊟ **Audio** | **6 g** | |
| 6 | translation | 1 g | |
| 7 | audio recording | 3 g | 6 |
| 8 | audio post-production | 2 g | 7 |
| 9 | ⊟ **Graphics** | **6 g** | |
| 10 | translation | 1 g | 13;15 |
| 11 | DTP | 5 g | 10 |
| 12 | ⊟ **Dynamic on screen text** | **1 g** | |
| 13 | translation | 1 g | 15 |
| 14 | ⊟ **Interactive elements** | **11 g** | |
| 15 | translation | 1 g | 2 |
| 16 | Software integration | 3 g | 15;13;1;5;9 |
| 17 | ⊟ **Audio redo** | **4 g** | |
| 18 | translation | 1 g | 14 |
| 19 | audio recording | 2 g | 18 |
| 20 | audio post-production | 1 g | 19 |
| 21 | **Software integration** | **1 g** | 17 |

Figure 6. schedule for localization process 2 (real)

The next section describes some practical examples of multimedia products for training and advertising that show the role of the *constraints* and that of the *driving localization components* in the localization process.

## *Examples*

### Online tutorial

The main goal of an online tutorial is to explain to the user a set of concepts through the commented visualization of simulated situations.

This format is very challenging as in one situation it is the audio script that contains the most important information, in a subsequent situation it is the screen that shows a specific content that can't be changed (the online services interface). Also, elements that appear on screen repeatedly: various elements (like buttons, menus) need to be represented in a consistent way throughout the presentation.
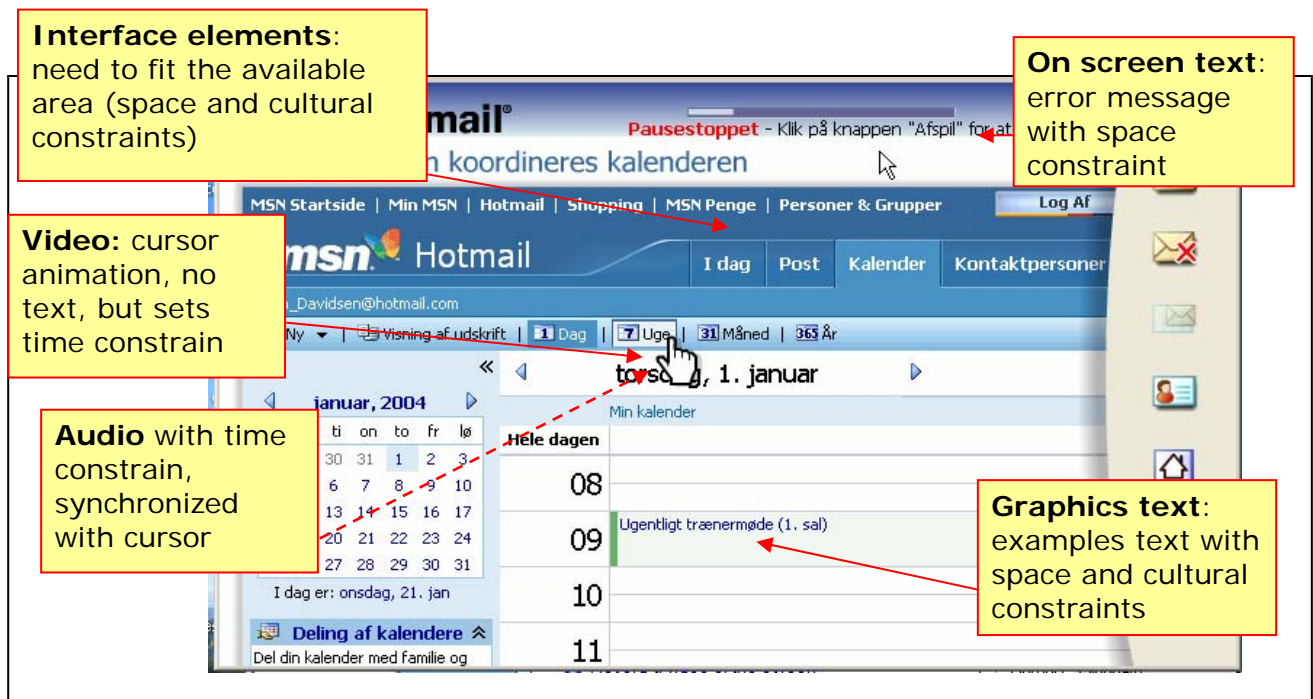


Figure 7. components and constraints in an online tutorial

Typical issues encountered with online tutorials are:
1. translated text for speech does not fit the time constraints, so that source code of the interactive tutorial must be modified,
2. screen shots do not fit the original size, so that the graphics layout has to be changed
3. translated text to appear on screen does not match the content of the audio.

It is evident that a process where all the text is translated first and then the rest of the process is adapted around it is not the shortest way to get the required result. The final QA process would outline several compatibility problems.

In order to get an effective localized product, we must:
- study the way the tutorial is built in terms of media and interactive software engine
- identify correlations (potential conflicts) between components
- be prepared to adapt the layout of the visual elements
- study the most efficient way to solve the conflicts by identifying the relevant constraints.
- be prepared to compromise whenever a conflict of constraints arises
- brief the translator about such constraints
- provide the translator the running version of the tutorial for context,
- brief the relevant team members (programmer, actors, sound and DTP engineers) about the constraints

Here is in this case the cross-component analysis:
- The action timing is to be retained at all time, to avoid complex software re-synchronization, different for each language (21 in this example)
- Time is a major constraint, that affects audio
- Audio text depends on examples
- Audio does not depend on error messages
- Audio depends on interface elements
- Audio depends on graphics
- Graphics: large amount of screens to be shot, and edited
- Graphics depends on examples
- Graphics does not depend on the on screen text
- Examples have space constraints, but have space constraint
- Interactive elements are drawn from the actual portal, and are already translated

In this case therefore the *driving localization components* are the graphic text and the audio. We need to translate the audio early so that when record it, the original length is preserved, to ensure that the new audio runs in synch with the rest of the presentation.
At the same time, the graphic production must start early as it is time consuming due to the size of the course. This in turn requires the interactive elements to be accessed on line (the course is about the features of a portal) before translating the other components, to be sure that cross-references are accounted for.
The correct localization process therefore is the one displayed in figure 8.
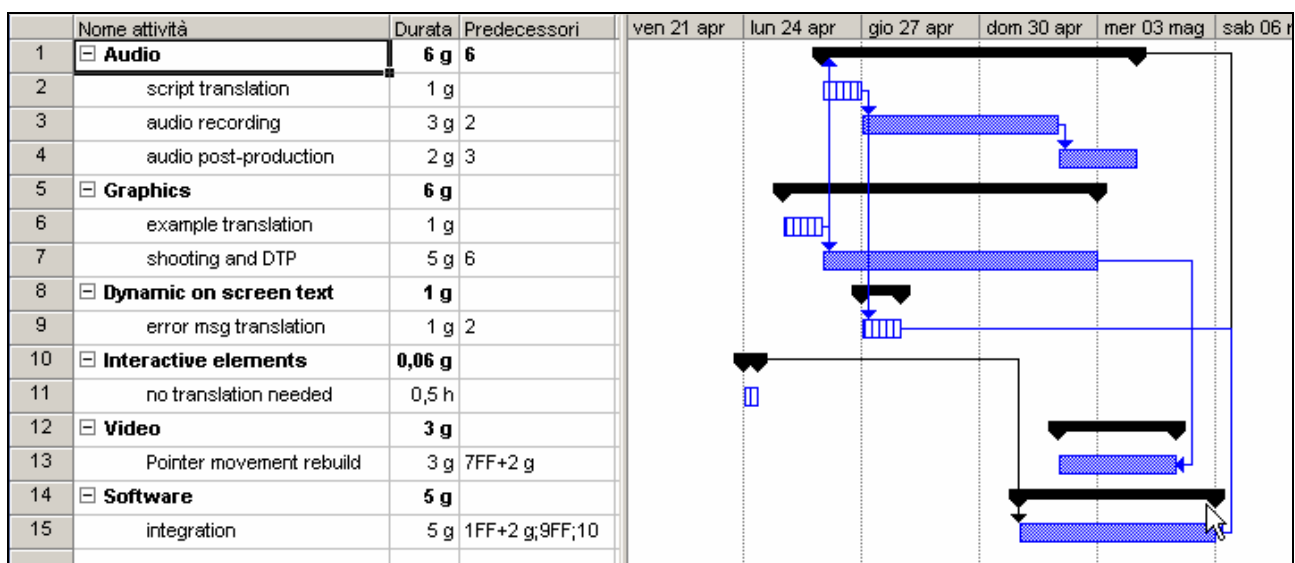


Figure 8. correct process for localization

## Web advert: IBM example

We recently worked on the launch of a new page on the IBM On Demand Business website in four European countries: Germany, Italy, France and the UK. The centre piece of the new content was a video about an IBM expert explaining the importance of innovation.

Unlike other pages from the same website, where an article would be the centre piece, here the video was, and this realization led to a series of changes to the design of the original American page for the European markets, as well as to a complete change in the prioritization of work.

On the American page, which was used as a source, there was first a landing page with a big picture of the expert and links leading to a second-level page where the video could be viewed on a video player or, alternatively, the PDF with the transcript could be read.

Since the video was the most important part of the communication, what was decided for Europe was - first of all - to move the video player to the landing page in a prominent position.

By applying the methodology of the constraint tables we then had to consider what the **driving components** of the localization would be.
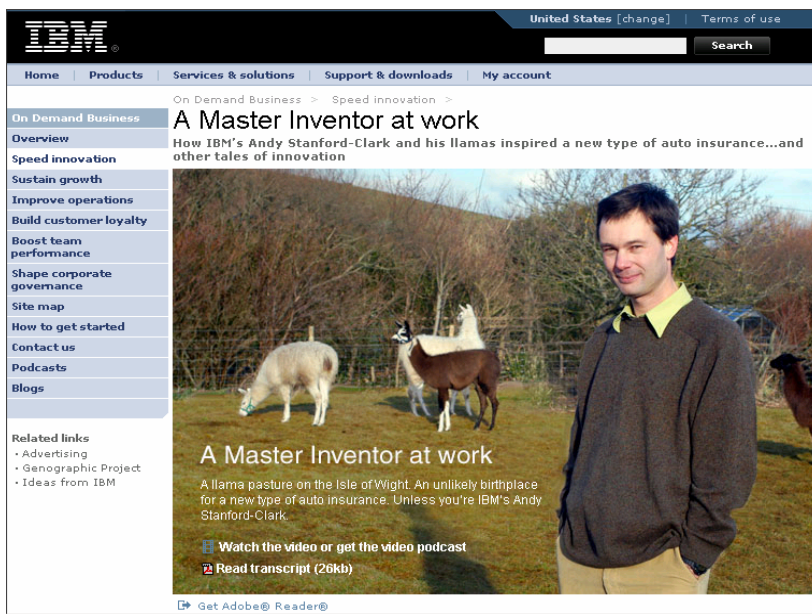


*US landing page. Video podcast only available through small link*

At the back of our minds we also had to remember to keep costs down so, at first, led by our old text-led methodology, we wanted to translate the transcript and use the translation to create subtitles. However we soon realized that this was an impractical approach. The speaker was extremely fast and subtitles would have had to run so fast that no one would have been able to read them, especially in languages that are longer than English. Alternatively they would have covered half the screen and given the fact that they were going to be displayed on a computer's video reader they would have been unreadable anyway.

The other consideration was that in countries such as France, Italy and Germany, consumers are not very accustomed to subtitles, and prefer dubbing or voiceovers.

We decided that the voiceover was the best option as it left some taste of the original in the background, but allowed to communicate directly to the target in their language.

As we are talking about voiceovers, the most significant constraint is time.

Although the translation of the video was very limited in size, the titles appearing before each section and the last frame with the American telephone numbers had to be reworked.as well. The

longer texts would have not allowed to recreate the same sort of animation out of readability problems, but also because the animation was embedded in the video and could no longer be replaced in foreign language versions.

Hence we concluded that the driving components of this project were
1. **the audio**, which was mandatory to translate according to the time constraints in order to be laid on the video.
2. **the video**, which had to be translated according to the space constraints given by the new font size selected.

The translation of the audio script was organized as follows.
Before starting the translation process, we created a table for the translation of the transcript in which we stated the maximum number of characters allowed for each paragraph based on the time needed to read them in each language and taking into account the little delay at the start of each section required by voiceovers. Translators had to adapt the text into their normally longer languages to fit a space actually smaller than the English original.
By doing so, we managed to avoid the need for extra adaptations and re-recordings which often mar work in text-led projects.

The recording was done in each of the target markets, for greater choice and quality of actors, and the post-production was centralized to reduce costs and retain a greater general overview of the project and achieve consistent quality of the audio.



*German page: video available on landing page and addition of chapters for easier use.*

The text for the video was extracted and inserted into a table in which we stated the maximum number of characters allowed for each frame and briefed the translator that the last frame, with references to phone numbers, had to be filled in with new data provided by IBM.
When we set foot in the video studio the whole chain, from the client to the translator were aware of the relevant constraints, and this enabled us to do the work in a single take, without any need for expensive redos.
The result was indeed very satisfactory and allowed IBM's European markets to enjoy a well-targeted piece of communication with an equal level of impact in each language.

The only point that could not be rendered as nicely as we would have liked - because the video was made before thinking of its localization - was the fact that the animated titles could not be recreated as they were embedded in the film and could nor be removed. The solution was to replace them with a still frame bearing the localized title, a so-called billboard. In this case we were lucky that the animation did not run on top of live footage, otherwise it would have been much more difficult to integrate billboards with the titles.
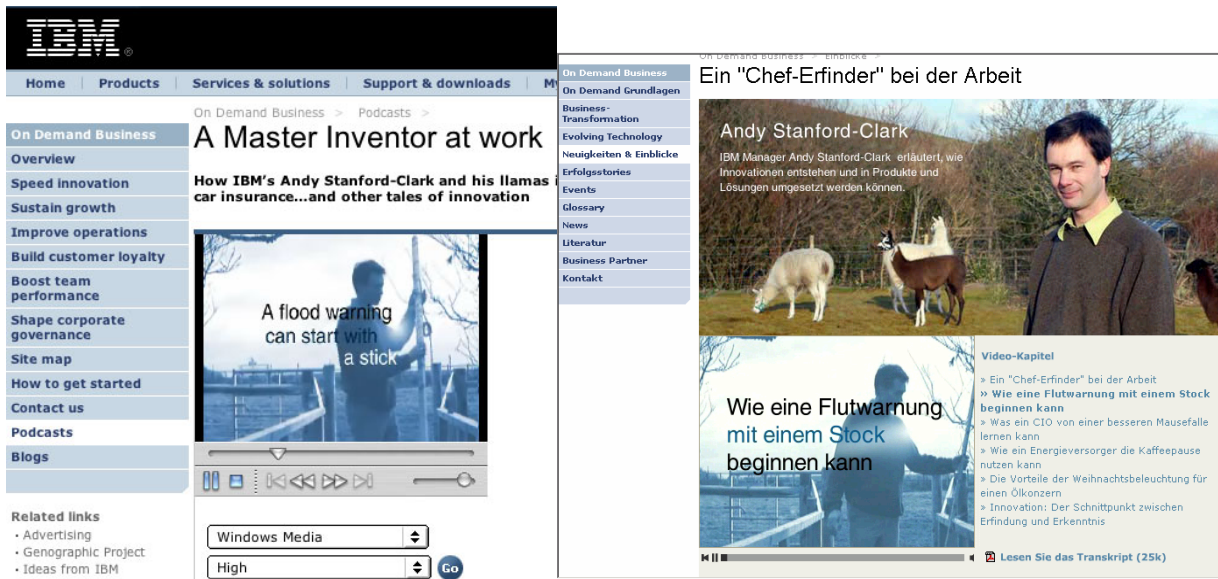


*Title localization: US animated title had to transform into German still frame.*


**When things go wrong: an example**
A real-life example we could make of things going wrong is that of a large corporation, which can't be named for legal reasons, which recently decided to launch a promotional campaign for a product across several countries, in which the TV campaign, the Direct Marketing activity and the online activity was each led by a different department inside that corporation, none communicating with each other. The client briefed different departments of the advertising agency, and each department had the same content translated separately. This meant that the campaign message was translated three times (loss of money) in three different ways (loss of consistency) and each without taking into account the constraints presented by the other media. What's more the TV ad and the DM materials were to drive traffic to both Points of Sale and to the website. And the website was to generate traffic to the Point of Sale. The DM was also responsible for setting up the POS. Each element was inextricably linked with the others, and the discrepancies were therefore going to become impossible to miss by the target audience. A few days before the launch of the campaign, by pure chance, a member of the online department noticed the discrepancy in the advertising headline on the TV ad's superscript, and the one used online in his native language. This prompted other checks, and the realisation of across-the-board inconsistencies in all languages. Urgent editing work had to be carried out at an enormous cost, including some re-work on the TV ad which actually cost less to change than remaking all the POS materials, and even so, in some cases the discrepancies could not be removed.

This case really highlighted how interdependent the various channels of communication were, and how a concerted approach was needed to help prioritize and plan localization activities in today's multimedia and multi-channel environment.

## *Conclusion*

The attempt by the multimedia localisation industry to find a systematic solution to the challenges posed by multiple constraints can be highly beneficial for global advertisers. The constraint tables help us understand that there is no longer one privileged starting point for the localisation process, and that both commercial communication and its localization have acquired a multidimensional nature. We now need to think in terms of priorities determined by driving components, and in order to do that the localization process of commercial communication has to start at campaign concept level.

For the translator, the challange is to work on texts that are deployed in a variegate set of components requiring a multi-dimensional approach to translation to fit in the localization process as a whole.

## *Bibliography*

- Szarkowska Agnieszka , "The Power of Film Translation", Translation Journal, Volume 9, No. 2 April 2005
- Savourel Yves, XML Internationalization and Localization, SAMS, 2001, pp. 126-133
- Esselink Bert, A practical guide to localization, J. Benjamins Publishing, 2000
- Walsh Aaron E., Bourges-Sevenier Mikaël , MPEG-4 Jump Start, Prentice-Hall PTR, 2000, Chapter 1
- Chapman Nigel, Chapman Jenny, Digital Multimedia, J Wiley ans Sons, 2000, pp 348-367
- Maroto Jesus, De Bortoli Mario, Languages in international advertising campaigns: the transcreator, International Journal of Localisation (2004). Localisation Research Centre. Limerick. Ireland.
- Maroto Jesus, De Bortoli Mario, Localisation and webvertising: conveying the digital
- communication mix, Proceedings of the eContent Localisation conference, 2002, University College Dublin, Ireland
- Maroto Jesus, De Bortoli Mario, Web site localisation, Proceedings of the European Languages and the Implementation of Communication and Information Technologies (Elicit) conference, 2001, University of Paisley, UK.
- De Bortoli Mario, Maroto Jesus, Translating colours in web site localisation, Proceedings of the European Languages and the Implementation of Communication and Information Technologies
- (Elicit) conference, 2001, University of Paisley, UK.