Translation Quality Measurement in Practice

Riccardo Schiaffino

President, Aliquantum

Franco Zearo

Worldwide Director, Language Services, L10nbridge Technologies

46th ATA Conference, Seattle

ALIQUANTUM

L10NBRIDGE

Introduction

- **When did you start working on translation quality measurement?**
  - We started thinking about the possibility (and necessity!) of measuring translation quality at the ATA conference in Orlando.
  - During the next three years we worked on developing a TQM strategy and tools.
  - The methodology is now in practical use at Lionbridge, and we keep working on improving it.
  - Aliquantum offers custom translation quality solutions.

ALIQUANTUM

LIONBRIDGE

Introduction

- **Why try to measure translation quality?**

   Without some means to assess the quality of translation, it is not possible to improve translation quality, nor is it possible to know if the translation quality is good; and, if it is good, how to keep it that way.

ALIQUANTUM

LioNBRIDGE

Introduction

- **Is it possible to measure translation quality?**

  We believe it is possible to measure translation quality, albeit indirectly:

  When measuring translation quality, we really measure the incidence of various types of errors and defects in the translated material

  - errors of meaning
  - errors of form
  - errors of compliance

  We define a good translation as one in which no (or few) errors are made.

ALIQUANTUM

LIoNBRIDGE

Introduction

- **Why measure errors when measuring translation quality?**

  The assessment of translation quality should be as objective as possible.

  We believe it is easier to agree on what constitutes an error rather than on what constitutes "quality" in the abstract, and that an important factor in quality is the absence of errors.

ALIQUANTUM

LIONBRIDGE

Introduction

- **Is there one "ideal" translation process to ensure the best possible quality level?**

  The important thing is that the process we use should yield the desired result.

  The very purpose of translation measurement is to obtain useful information for benchmarking the relative merits of various translation processes.

ALIQUANTUM

LIONBRIDGE

- **How does translation quality measurement differ from other methods of translation quality assessment?**

  Many methods of evaluating translation quality have been developed and proposed.

  Malcom Williams classifies these methods into two categories:

  - Argumentation-centered systems
  - Quantitative-centered systems

Introduction

- **Argumentation-centered systems**

  More holistic approach

- **Quantitative-centered systems**

  Characterized by some method of error counting

  The advantage of quantitative-centered methods is that they lend themselves to quantifying errors and, therefore, make measurements possible.

ALIQUANTUM

LIONBRIDGE

The Translation Quality Index (TQI)

- **What is the TQI methodology?**

  The TQI methodology is a quantitative-based method of translation quality assessment.

  It measures the number and type of errors found in a text and calculates a score, or TQI, which is indicative of the quality of a given translation.

ALIQUANTUM

LioNBRIDGE

The Translation Quality Index (TQI)

- **Translation Quality Index**
  - A number that is indicative of the quality of a given translation sample.
  - Obtained by the rigorous application of a quality assurance methodology.

  The Translation Quality Index attributes a value to a translated text, with 100 being an "error-free" translation. It is based on the number of error points in a given text or sample.

ALIQUANTUM

LioNBRIDGE

The Translation Quality Index (TQI)

- **What are error points, and how do they differ from errors?**

  Not all errors are equal: There is a difference between a typo on the front cover of a manual and the same typo in a footnote.

  This leads us to assign different weights to errors depending on their consequences. In our previous example, we can decide to give minor typos a weight of "1", and major typos a greater weight.

  We call these weights "error points".

ALIQUANTUM

LioNBRIDGE

The Translation Quality Index (TQI)

- **Separation between error type and severity**

  No pre-assigned penalties for the different error categories.

  Each error can be marked as critical, major, or minor, depending on its consequences.

The Translation Quality Index (TQI)

- **Strict criteria for the severity levels of errors**

  A TQI measurement should be objective, reproducible, and repeatable.

  To achieve these criteria, the evaluator has to follow certain rules when marking errors.

ALIQUANTUM

LioNBRIDGE

The Translation Quality Index (TQI)

- **What were the difficulties when you started to put the TQI into practice?**

  - The tendency to fit the data to the evaluator's quick perception of how good or bad a particular translation sample seems to be

  - Accuracy errors are difficult to evaluate when there is a slight loss in meaning

  - Even grammatical errors are sometimes not as straightforward as one would think

ALIQUANTUM

LIoNBRIDGE

The Translation Quality Index (TQI)

- **What makes a good evaluator?**
    - Must be able to be as objective as possible
    - Must be able to distinguish between factual, tangible errors and stylistic preferences

  What would be helpful is a certification program for evaluators, sponsored by an independent, not-for-profit organization such as the ATA.

  This not-for-profit organization might create standards regarding error classification, severity levels, error points, and others.

ALIQUANTUM

LioNBRIDGE

The Translation Quality Index (TQI)

- **How do you distinguish between errors and stylistic preferences?**

    Stylistic preferences are not errors and are ignored in the computation of the quality score. Therefore, it is necessary to establish clear rules that define what is an error and what is not an error.

    The evaluator has to answer the following three questions:

    - Is it grammatically correct?
    - Is the translation accurate?
    - Is the translation compliant with the glossary, style guide, guidelines, and client instructions?

    If the answer to these questions is "Yes", it means that it is not an error.

ALIQUANTUM

LIONBRIDGE

The Translation Quality Index (TQI)

- **Is there anything that the TQI methodology cannot measure?**

  The TQI methodology is designed to measure tangible, factual errors only, it tends to be ineffective is when a high degree of creativity is expected on the translator's part, which is often the case, for example, with translations for marketing and advertising. In these types of text, translators and copyeditors might have a certain degree of freedom. It is an acceptable practice to deviate from the source text as long as the translator maintains the core message.

ALIQUANTUM

LioNBRIDGE

Translation Error Log Program

Purpose of this program is to aid the test reviewer to record in a separate file any translation error found, together with the appropriate assessment values and any remark.

The output file will contain the data necessary to calculate statistic on translation quality as well as the TQI.

ALIQUANTUM

LioNBRIDGE

Translation Error Log Program

Instructions

1 - Assign an error category to the error, and then the weight
2 - When an error is ambiguous, always choose the earliest error category (e.g., "Error of meaning" before "Error of form")
3 - When in doubt, always choose "serious" over "minor"
4 - Indicate in the Remarks what the error is and what the correct translation should be

ALIQUANTUM

LioNBRIDGE

Translation Error Log Program

Error Category
- ○ Error of Meaning
- ○ Error of Form
- ○ Error of Compliance

Weight
- ○ Serious
- ○ Minor

Translation Error Log Program

- **Error Category**

   The type of error; errors should be categorized as:

    - Errors of Meaning
    - Errors of Form
    - Errors of Compliance

Translation Error Log Program

- ## Errors of Meaning
  - If the meaning of the translation is different than the meaning of the SL

- ## Errors of Form
  - If the translations contains an error of grammar, spelling, or other formal error that does not otherwise change the meaning of the translation with respect to the SL

- ## Errors of Compliance
  - If the translation, even though meaning and form are correct, does not conform to the instructions received, style guide, preferred terminology, or other customer-specified requirement.

ALIQUANTUM

LioNBRIDGE

Translation Error Log Program

- **Error Weight**
  - The gravity of the error. Errors should be categorized as either "serious" or "minor".
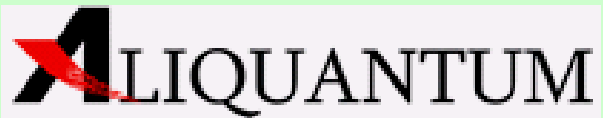
Translation Error Log Program

- Serious

  - Errors of meaning: The meaning of the TL is completely different than the meaning of the SL
  - Errors of form: A blatant error that any educated speaker of the TL would recognize as such
  - Errors of compliance: Blatant disregard of the instructions, style guide, etc. received

In all cases, an error that is in a very prominent position should be classified as serious.

TranslationQuality.com

ALIQUANTUM

LIoNBRIDGE

Translation Error Log Program

- Minor
  - Errors of meaning: The meaning of the TL is slightly different than the meaning of the SL.
  - Errors of form: A clear error concerning a minor point of grammar, punctuation, etc.
  - Errors of compliance: A minor deviation from the instructions, style guide, etc.

Translation Error Log Program

Translation Error Log - Help

Translation Error Log Help

Purpose of this program is to aid the test reviewer to record in a separate file any translation error found, together with the appropriate assessment values and any remark.

The log file will contain the data necessary to calculate statistic on translation quality.

Error Category

The type of error; errors should be categorized as errors of Meaning, errors of Form or errors of Compliance

Errors of Meaning
If the meaning of the translation is different than the meaning of the SL

Errors of Form
If the translations contains an error of grammar, spelling, or other formal error that does not otherwise change the meaning of the translation with respect to the SL

Errors of Compliance
If the translation, even though meaning and form are correct, does not conform to the instructions received, style guide, preferred terminology, or other customer-specified requirement.

Error Weight

The gravity of the error. Errors should be categorized as either "serious" or "minor".

Serious
In the case of errors of meaning: the meaning of the TL is completely different than the meaning of the SL - OR - an error that would otherwise be classified as minor, but in a very prominent position. In the case of errors of form, a blatant error that any educated speaker of the TL would recognize as such - OR - an error that would otherwise be classified as minor, but in a very prominent position. In the case of errors of compliance, blatant disregard of the instructions, style guide, etc. received - OR - an error that would otherwise be classified as minor, but in a very prominent position.

Minor
In the case of errors of meaning: the meaning of the TL is slightly different than the meaning of the SL.
In the case of errors of form, a clear error concerning a minor point of grammar, punctuation, etc. In the case of errors of compliance, a minor deviation from the instructions, style guide, etc. received

About

ALIQUANTUM

LIONBRIDGE

TQI Spreadsheet

ALIQUANTUM

Translation Quality Index Form

Summary

Review Details

ALIQUANTUM

LIONBRIDGE

TQI Spreadsheet

ALIQUANTUM

Language Quality Index Form

Review Details

	Language Name	Language Code	Date
Source language:	Italian	IT_IT	9-Nov-05
Target language:	English (USA)	EN_US	
Project:	Test-12-abcx	Verifier	
Number of words in sample:	250		R. Schiaffino

Location	Error Category	Severity	Source te
File test 1	Form	Minor	

ALIQUANTUM

LioNBRIDGE

TQI Spreadsheet

ALIQUANTUM

Translation Quality Index Form

Review Results

Source Language	Target Language	Verifier:	Date:
Italian	English (USA)	R. Schiaffino	9-Nov-05

	TQI	92
Result:	Pass	

Project	TEST-12-ABCX
Number of Words in Sample	250
Minimum TQI Allowed	90

Error Category	Minor errors	Major errors	Category Total (EP)
Form	1	0	1
Meaning	0	0	0
Compliance	0	0	0
Total	1	0	1

ALIQUANTUM

LioNBRIDGE

Customer-Driven Approach to Quality

Functional approach to quality:

- Meeting the needs and expectations of the customer or user

- Quality as a competitive weapon

- Improving on quality can reduce costs and speed up time-to-market

ALIQUANTUM

LioNBRIDGE

Why is Quality Measurement Important?

- It is difficult to improve something if you cannot measure it
- Metrics provide a way to objectively quantify a process
- Reduce the cost of poor quality
- Increase customer satisfaction
- Benchmarking
- Competitive advantages

ALIQUANTUM

LioNBRIDGE

How To Set Up a Language Quality Measurement System (1)

The six phases for setting up a robust quality system:

- Design
- Calibration
- Sampling
- Measurement
- Statistical Analysis
- Process Improvement

ALIQUANTUM

LioNBRIDGE

How To Set Up a Language Quality Measurement System (2)

Phase: 1 – Design

- Collect a corpus of good and bad translations
- Analyze the corpus to identify critical-to-quality errors (**error definition**)
- Decide what to measure (**error categorization**)
- Assign a weight to various types of errors (**severities**)
- Define an error threshold

Phase: 2 – Calibration

- Pilot, test, and adjust until the system works in an objective, repeatable, and reproducible way

ALIQUANTUM

LioNBRIDGE

How To Set Up a Language Quality Measurement System (3)

Phase: 3 – Sampling

- Sample selection criteria (e.g. random, systematic)
- Sample size, confidence intervals, margins of error
- Cost considerations (find the point of diminishing returns)

Phase: 4 – Measurement

- Calculation of a Translation Quality Index (TQI)
- Evaluation must be repeatable, reproducible, objective
- Use of independent auditors
- Provide formal training

ALIQUANTUM

LIONBRIDGE

How To Set Up a Language Quality Measurement System (4)

Phase: 5 – Analysis

- Use of control charts
- Investigate results
- Rule out special causes

ko_KR, May 2005
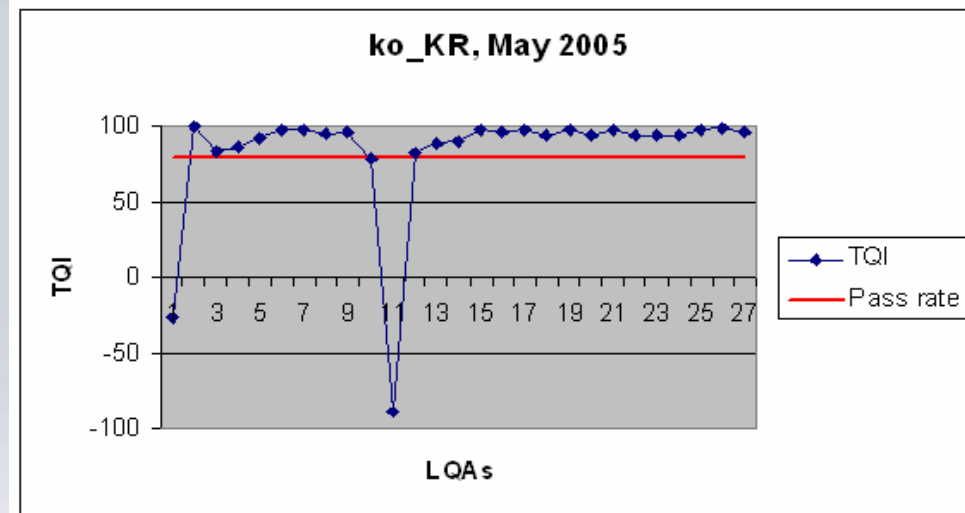
Phase: 6 – Process Improvement

- Take corrective actions
- Compare the TQI values before and after a process change to check for actual process improvement

ALIQUANTUM

LIoNBRIDGE

Language Quality Assurance Process (1)

- The LQA process provides a **quantifiable** means of measuring the quality of the translations that we receive from our translation partners.

- The LQA process involves auditing a **sample** of the files, usually at an early stage of the localization process.

- The LQA reports provide valuable information to **monitor the performance** of our translation partners.

- The LQA provides important data for continuous **process improvement**.

ALIQUANTUM

LioNBRIDGE

Language Quality Assurance Process (2)

- Trained LQA specialists look at the following error categories:
  - **Accuracy**
  - **Terminology**
  - **Language**
  - **Style**
  - **Country**
  - **Formatting**
  - (We can also include client-specific error categories)
- **Each error is assigned one of these severity levels:**
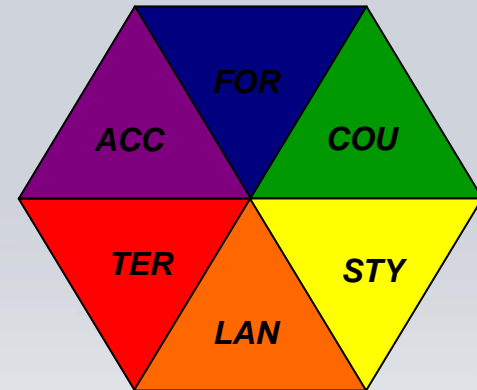
CRITICAL   MAJOR   MINOR   Preferential

ALIQUANTUM   LioNBRIDGE

Language Quality Assurance Process (3)

- The LQA process also provides a measure of translation quality known as the **Translation Quality Index** (TQI).
- The quality index is based on the number of errors and type of errors detected in a given sample.
- The quality index is like taking the "temperature" of the quality of a translation.
- The greater the TQI score, the better the translation.

100 90 80 70 60 50 40 30 20 10 0

ALIQUANTUM

LIONBRIDGE

Measuring Translation Quality

"Horrible"

"Poor"

"So-so…"

"Okay"

"Good"

Percentage of errors vs total word count

100% 90% 80% 70% 60% 50% 40% 30% 20% 10% 0%

5%
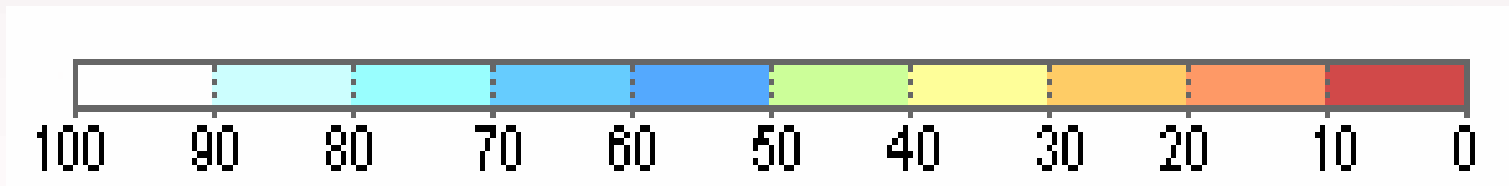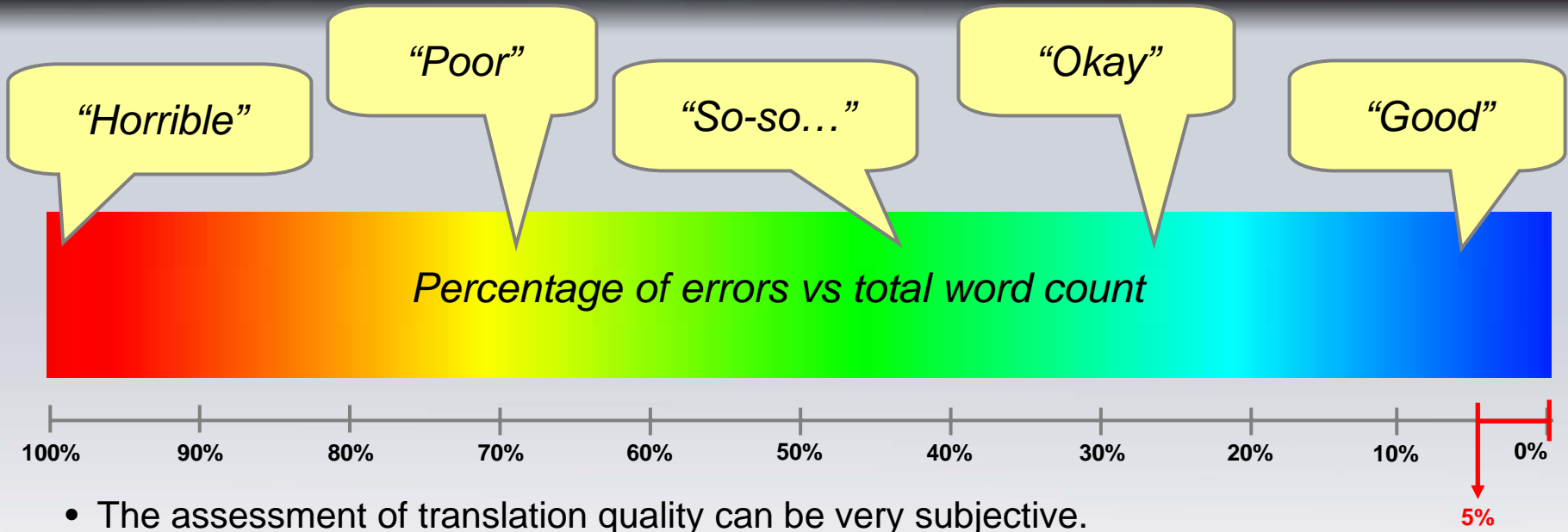
- The assessment of translation quality can be very subjective.

- A quantifiable means of measuring translation quality can help to reduce the subjectivity.

- The error point value of translations done by our professional translators very rarely exceeds 5%.

- We focus our attention on monitoring and improving this 5%.

ALIQUANTUM

LIONBRIDGE

Measuring Language Quality At Lionbridge

Lionbridge's Language Quality Assurance (LQA) process provides a **quantifiable** means of measuring the quality of the translations

We use our LQA form and tools to calculate a translation quality index (TQI)

The index attributes a weighted score to a sample taken from a translated text.

5%

Percentage of errors vs. total word count

TQI=100

TQI

TQI=0

A LIQUANTUM

LioNBRIDGE
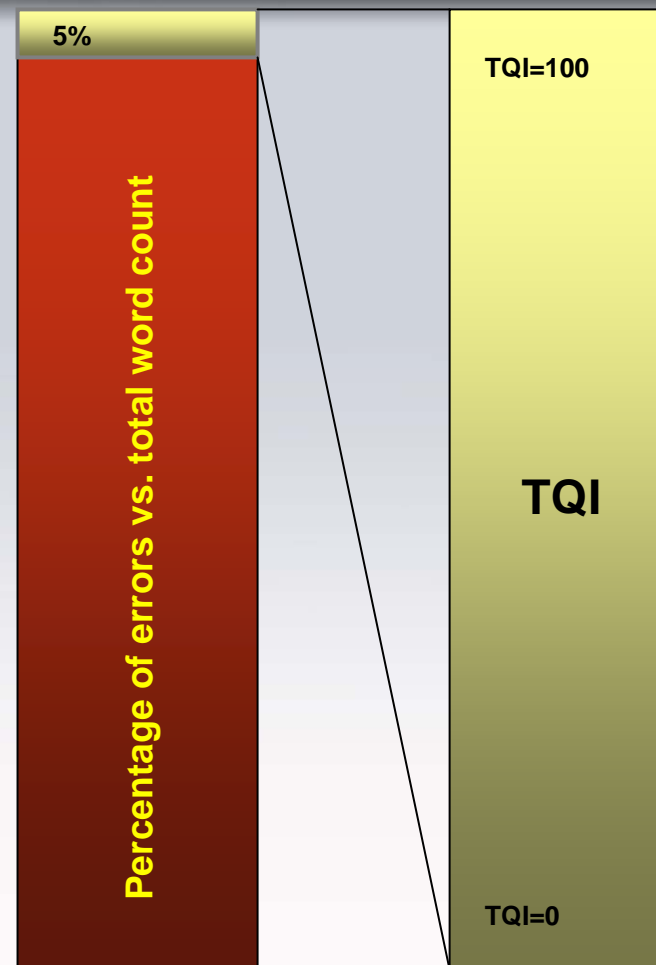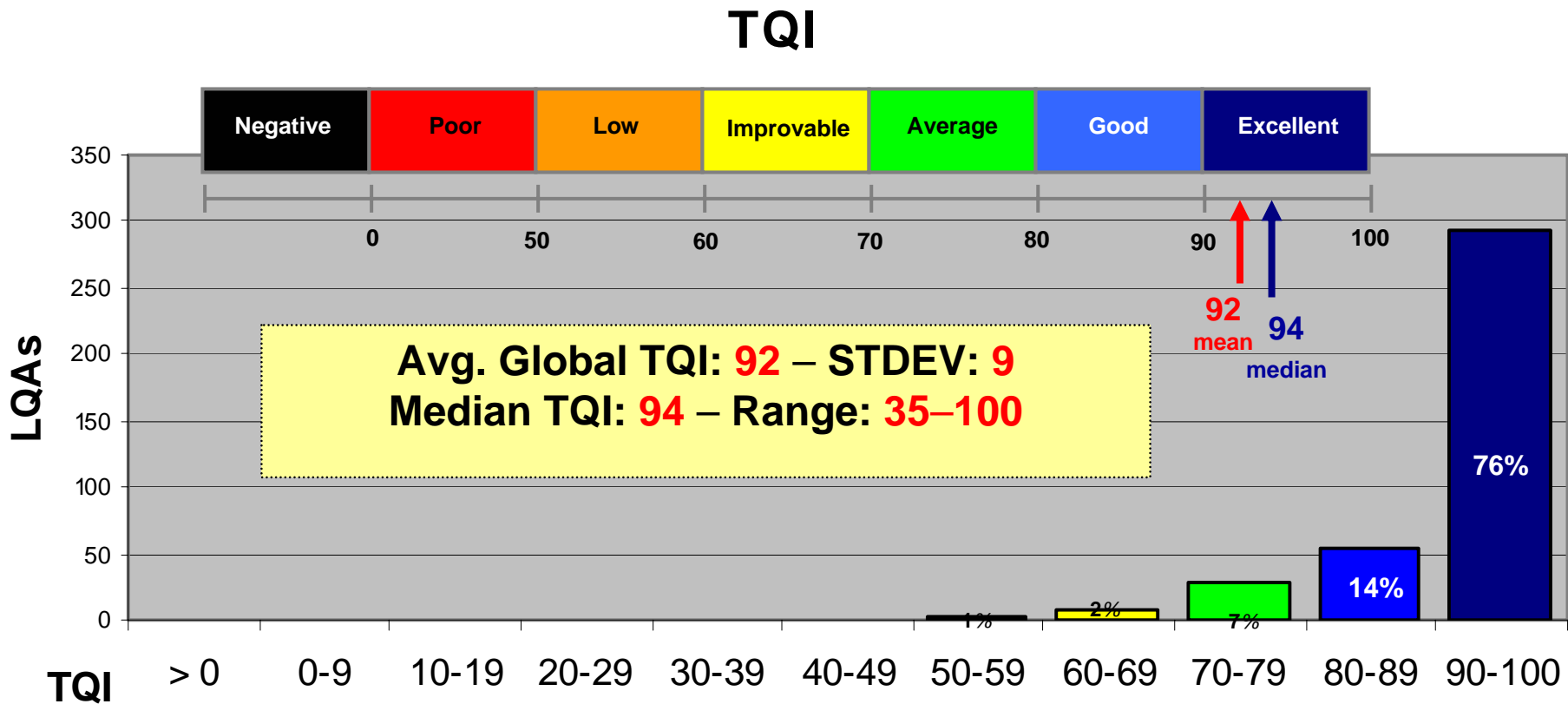
Case Study 1: Early Detection of Quality Problems

Project Summary

Project: **BELL4**
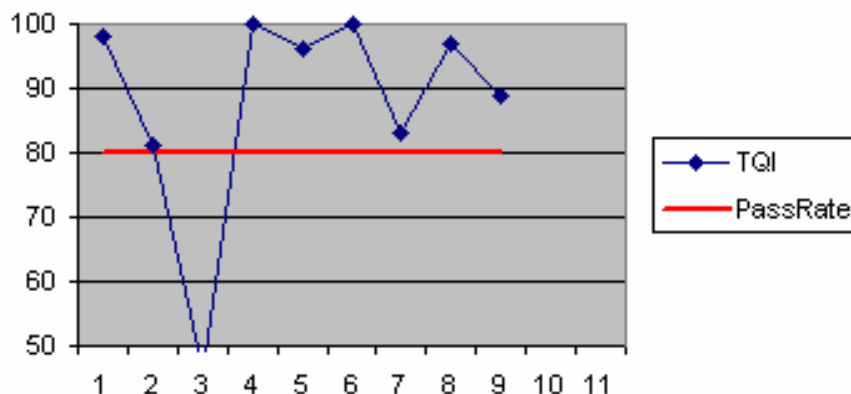
Average TQI: **88**

Project Number	Project Name	Client Name	Vendor name	Language Lead	LQA Specialist	Import Date	QA Date	Language Pair	Sample Size	TQI	Pass/Fail
ABCBD0403	BELL4	ACME Inc.	Translator1	Yasuyo Kitar	Yasuyo Kitar	9/14/2004	8/27/2004	EN-JA	1984	98	Pass
ABCBD0403	BELL4	ACME Inc.	Translator2	Carlos Lopez	Carlos Lopez	9/14/2004	8/27/2004	EN-ES	5342	81	Pass
ABCBD0403	BELL4	ACME Inc.	Translator3	Silvio Picinini	Silvio Picinini	9/14/2004	8/27/2004	EN-BR	1984	46	Fail
ABCBD0403	BELL4	ACME Inc.	Translator3	Silvio Picinini	Silvio Picinini	9/14/2004	8/27/2004	EN-BR	582	100	Pass
ABCBD0403	BELL4	ACME Inc.	Translator1	Yasuyo Kitar	Yasuyo Kitar	9/14/2004	9/9/2004	EN-JA	1106	96	Pass
ABCBD0403	BELL4	ACME Inc.	Translator2	Carlos Lopez	Carlos Lopez	9/14/2004	9/10/2004	EN-ES	1331	100	Pass
ABCBD0403	BELL4	ACME Inc.	Translator1	Yasuyo Kitar	Yasuyo Kitar	9/14/2004	9/13/2004	EN-JA	713	83	Pass
ABCBD0403	BELL4	ACME Inc.	Translator2	Carlos Lopez	Carlos Lopez	9/14/2004	9/13/2004	EN-ES	1293	97	Pass
ABCBD0403	BELL4	ACME Inc.	Translator3	Silvio Picinini	Silvio Picinini	9/14/2004	9/13/2004	EN-BR	713	89	Pass

ALIQUANTUM

LIONBRIDGE

Case Study 2: Errors vs. Preferential Variations

Language Quality Assurance Form

Review Results

Source Language	Target Language	Language Lead:	Date:	TQI	99
English (USA)	Japanese	Yasuyo Kitano	27-Aug-04		

Project Code	ABCBD04030
Project Name	BELL4
Component	online help
Number of Words in Sample	1984
Maximum Error Points (EP) Allowed	10

Result: Pass

Error Category	Minor errors	Major errors	Critical errors	Preferential	Category Total (EP)	Category Limit (EP)	Category Result	Areas of Attention
Accuracy	0	0	0	0	0	2		
Terminology	0	0	0	0	0	2		
Language Quality	0	0	0	0	0	2		
Style Guide	1	0	0	0	1	2		
Country Standards	0	0	0	18	0	2		
Formatting	0	0	0	0	0	2		
Client Specific	0	0	0	0	0	2		
Total	1	0	0	18	1			

LIONBRIDGE

ALIQUANTUM

LIONBRIDGE

Case Study 3: Korean Corrective Actions

ko_KR, May 2005

LQAs

TQI

Pass rate

ALIQUANTUM

LioNBRIDGE
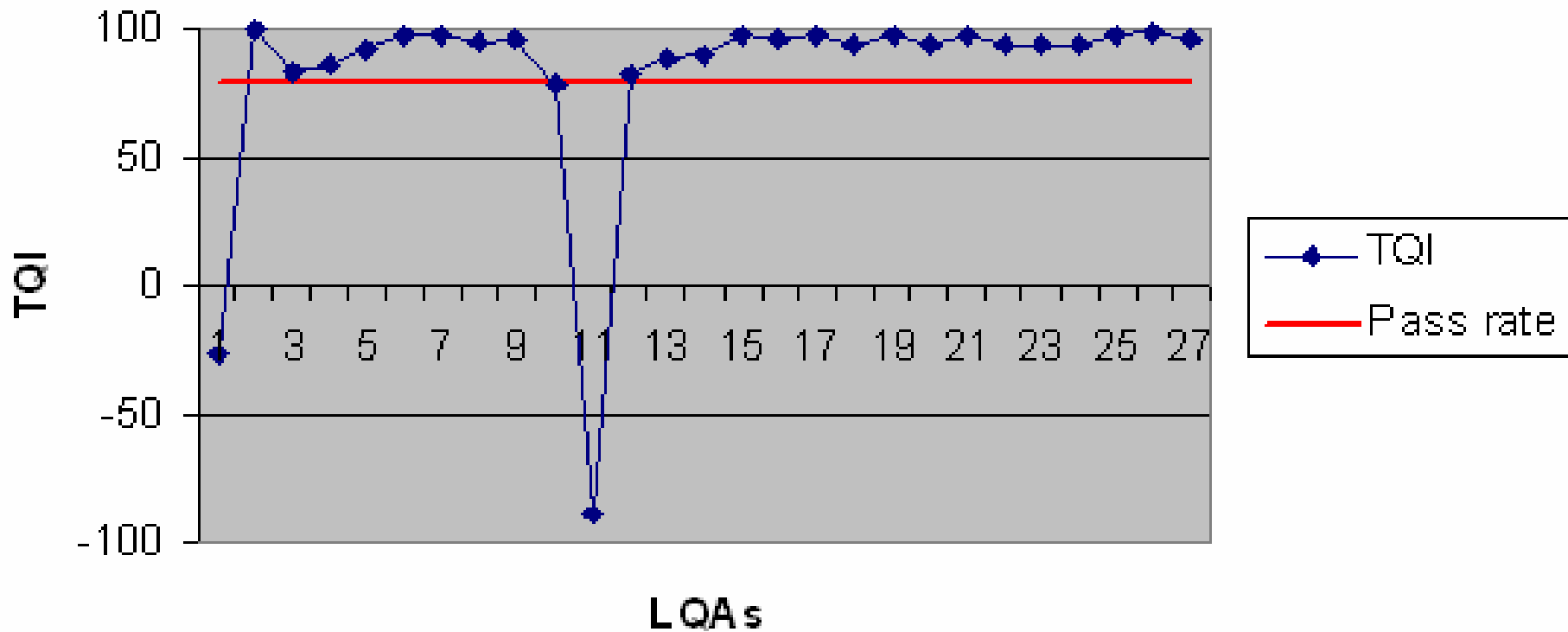
Case Study 4: Measuring Process Improvement
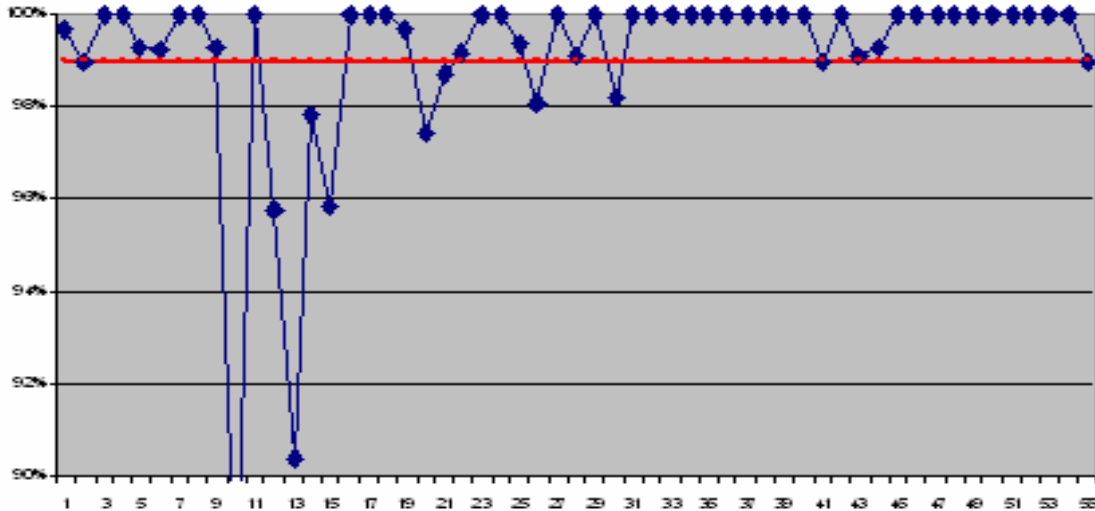
August 2005

Case Study 4: Measuring Process Improvement

August 2005

September 2005
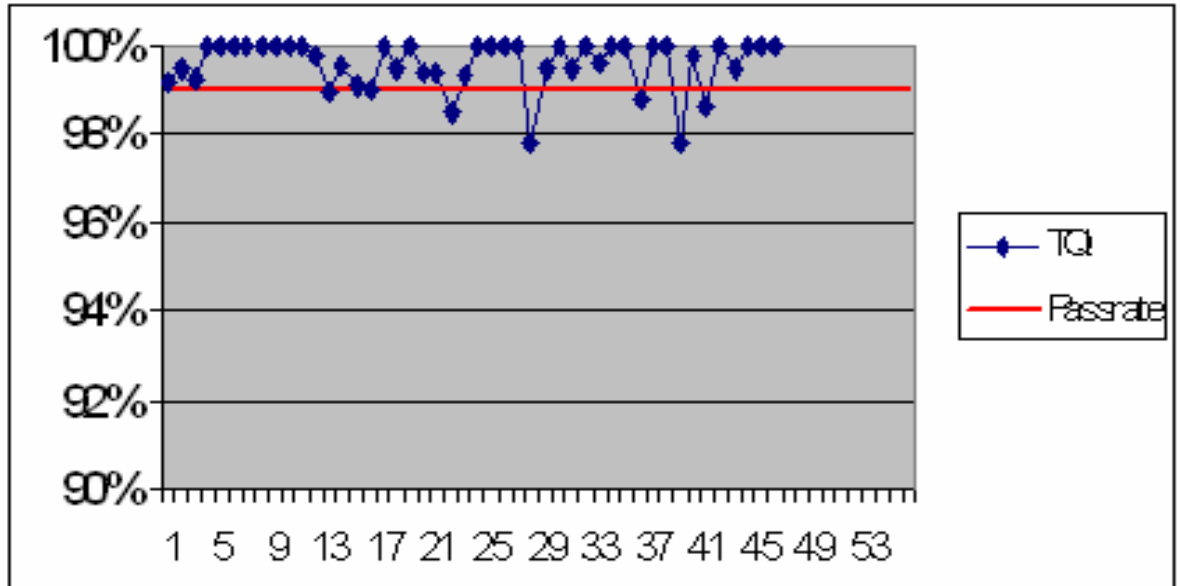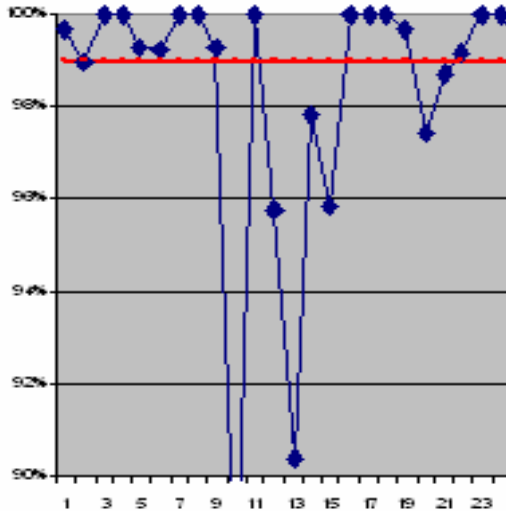
Caveats

Averages can be calculated to nineteen places of decimal with astonishing ease.

When the job is done, it looks very accurate.

It is an easy and fatal step to think that the accuracy of our arithmetic is equivalent to the accuracy of our knowledge about the problem in hand.

•M.J. Moroney, *Facts from Figures*

Although the use of points may impart a certain impression of objectivity, it is in truth still subjective.

•*ATA Reminder*

This is true, but our objective in developing our system is precisely to increase objectivity in the evaluation of translations

ALIQUANTUM

LIONBRIDGE

Limitations

- **What are the limitations of the TQI methodology?**

  - Not suitable for creative texts (e.g., marketing)
  - Results may be unreliable if samples are too small
  - Results are meaningless if not enough samples are taken
  - Risk of of setting up an overly complicated system

  Measurement is only **one** component of an entire system

ALIQUANTUM

LIoNBRIDGE

Thank You!

Additional Resources

Copies of this and of our previous presentations, a translation quality blog, and other materials can be found on our website, www.translationquality.com.

Riccardo_Schiaffino@aliquantum.biz
Franco.Zearo@lionbridge.com

ALIQUANTUM

LioNBRIDGE