Simultaneous Interpretation of Numbers: Comparing German and English to Italian. An Experimental Study

Diletta Pinochi

Free-lance Conference Interpreter

Abstract

An experimental study was carried out to investigate whether the difficulty of delivering numbers in SI is language-independent or whether some specific features – such as the different structures of the numerical systems in SL and TL – may also be relevant and influence SI performance negatively.

To this end, a German text and an English text, both dense with numbers, were interpreted simultaneously into Italian by 16 students. The first language pair (EN-IT) had a linear numerical system and the second one (DE-IT) did not, as in German the so-called inversion rule has to be applied.

An initial analysis of the results suggested that the difficulty of delivering numbers in SI is language-independent. However, a more detailed analysis of the outcomes showed that a significant difference between the two language pairs was apparent in the distribution and typology of errors: transposition/position errors (including inversion errors) were evident in German but not in English.

1. Introduction

In the existing literature on interpretation, numbers are often referred to as one of the most common "problem triggers", yet only a few experimental studies[1] have focused on this issue in an attempt to

1 In Italy: Alessandrini (1990), Crevatin (1990), Braun and Clarici (1997), Mazza (2000).

investigate the causes of such difficulties or propose any solutions to overcome them. Moreover, very different methodologies have been used in the analysis of these studies' results, which have also been based exclusively on just one language pair each time,[2] thus not allowing for a cross-linguistic detailed comparison between multiple language pairs.

This discussion document aims to provide a cross-linguistic analysis on the performance of simultaneous interpretation (SI) of a speech that is dense with numbers in two language pairs: German-Italian and English-Italian.

The final objective is twofold: on the one hand to shed light on the *universal causes* which underlie the (often) high percentage of errors made in the SI of a speech dense with numbers; on the other hand to investigate if some *particular causes*, connected to language specific structures, also play an important role. More precisely, the aim is to find out whether, in addition to *language-independent* causes of difficulty in the SI of a speech dense with numbers, there may also be some *language-dependent* factors and if so, to what extent, in order to determine whether the *language factor* is relevant.

The language pairs chosen for this study are characterized by a different numerical structure. In the first pair (English-Italian) the structural correspondence of the digits order is quite linear, and in the other one (German-Italian) it is not, as in German the so-called inversion rule is applied.

2. The numerical system: an overview

When dealing with numerical systems it is necessary first to distinguish between "numbers" and "numerals". Numbers are arithmetical objects, whereas numerals are the names used to name them (Hurford 1987).

Each number can be expressed through two different modalities, written and oral. These, in turn, can be represented through at least three different codes in all (Deloche and Seron 1987): the oral, phonological verbal code [/faiv/], the written graphemic or alphabetical code [*five*] and the written Arabic code [5].

Like all other linguistic systems, the numerical system also has its own lexicon/vocabulary (digits in the case of Arabic numbers and numerals in the case of verbal codes), its own syntax, which determines and regulates the relationship between the digits or the numerals, and its own semantic dimension. It is worth noting how the oral numerical system sometimes differs from the written one and how the verbal numerical systems of the languages analysed in this study differ from one another.

2 Italian-German (Braun and Clarici 1997) and Italian-English (Alessandrini 1990, Crevatin 1990, Mazza 2000).

The Arabic code is without doubt the only system common to the languages in this study. The universality of this code is opposed to the specific features of the various verbal numerals in any particular language. In SI, where all stimuli are perceived through the acoustic channel, one obviously has to deal with numerals. But that does not mean that numbers are extraneous to the interpretation task, as most interpreters and interpretation students (as this study aims to confirm) tend to note down numbers in the Arabic code (i.e. 5) when they are heard. In doing so, the interpreters can detach themselves from the phonological surface of the source language (SL) and of the target language (TL) using a neutral, "visual" representation of the number to be interpreted.

The code most relevant to the present analysis, however, is the oral phonological verbal code, i.e. numerals, which will be briefly described in the following sections.

2.1 The verbal numerical code

Deloche and Seron (1982, 1987) gathered important information about the syntactic and lexical mechanisms regulating the verbal numerical system. They analysed the errors made by brain-damaged patients writing down numbers to dictation and reading them aloud with the aim of using the results to describe the normal functioning of the mechanism.

Their hypothesis was that the numerical lexical system is composed of primitive elements and miscellaneous elements. The set of primitive elements is made up of units, teens and tens, while the miscellaneous elements are the multiplicators (hundred, thousand etc.).

On the basis of the actual positioning of the miscellaneous elements, the above-mentioned components combine through syntactic mechanisms and form complex denominations through additive or multiplicative relations (i.e. 300= 3x100; 103= 100+3). In this way it is possible to create infinite denominations starting from a limited set of items.

The linguistic formulation of the numerical relations upon which the construction process of numerals is based can be represented as follows:

- Lexicalization (i.e. a simple, new word: cinque, five, fünf);
- Addition (i.e. a complex numeral obtained by the addition of its elements: trentadue, thirty-two, zweiunddreißig > 30+2);
- Multiplication (i.e. a complex numeral obtained by the multiplication of its elements: ottocento, eight hundred, achthundert > 8x100).

The system is thus a hybrid one, well exemplified by the languages analysed in this study, which can be illustrated by a tree structure as shown in figure 1 below (350.272):

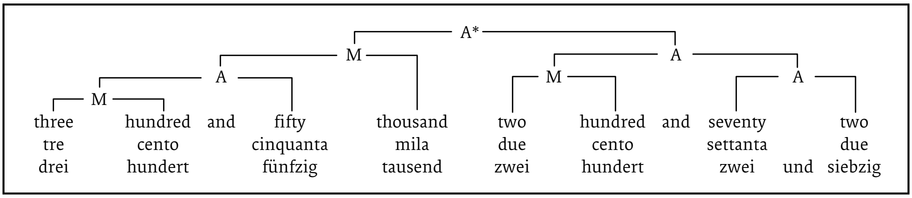

Figure 1.

So far, the three systems analysed – Italian, German and English – seem to follow the same construction rules. However, the German numerical system differs considerably from the English and the Italian ones, as described in the following section.

2.2 The German numerical system

Like the Italian and English numerical systems, the German system is constructed according to the rules described in the previous paragraphs. However, the German system has a major difference, that of non-linearity between the Arabic and the verbal code.

For instance, the Arabic code is universally visually understood and read from left to right, whereas some German numbers are pronounced from right to left. This requires the application of the "inversion rule", according to which "25" will be pronounced "five and twenty*" instead of "twenty-five".

This rule is applied to:

- All numerals with the ending "–zehn" (drei-zehn, vier-zehn...);
- All groups composed by teens + units (ein-und-zwanzig, zwei-und-zwanzig...), except when the unit is equivalent to multiples of tens (zehn, zwanzig, dreißig...).

Ths means that the inversion rule applies to all numbers between 13 and 99 (except for the tens), i.e. 79 numerals out of 100.

In transcoding a German number, i.e. switching from one code to another – from the Arabic to the verbal code or vice-versa – the processing of the number is not linear and requires one the performance of a series of non-linear, energy-consuming operations as shown in figure 2 below (Bosshardt 2004), regarding the number 32, 528, 331.

* A= Addition; M= Multiplication.
Representation by Dehaene 1992, slightly modified to be adapted to the present study.

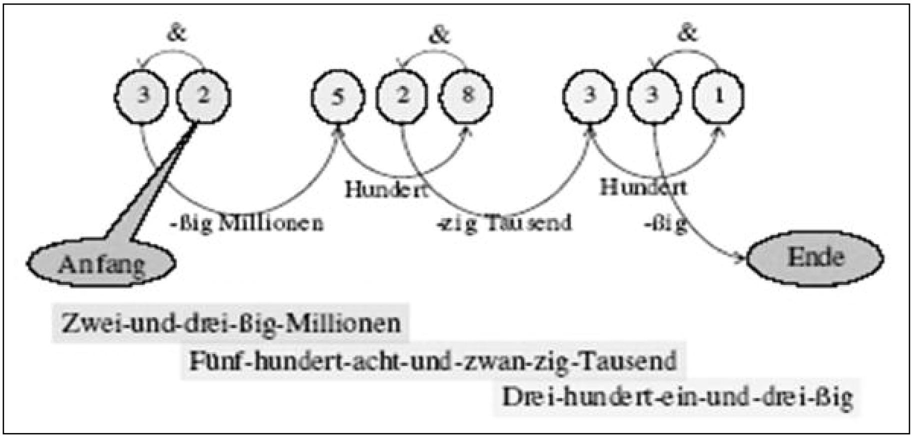

Figure 2.

2.3 The "Verein Zwanzigeins"

Starting from the statement made in 1915 that:

> Bei jedem Übertragen von Zahlwörtern in Ziffern, oder beim Lesen von Zahlen [...] muss man eine gewisse Denkarbeit für die Umstellung aufwenden, die das Behalten der Zahlen erschwert und häufig Veranlassung zu Fehlern gibt

the Mathematics and Psychology Department of the Bochum University in Germany launched a proposal in 2004 which led to the foundation of an association with the very apt name of "Verein Zwanzigeins".[3]

With regard to the "opposite verbalization" of German numbers, the association aims to change number reading in German in accordance with the linearity of the Arabic code. This would mean that "einundzwanzig" would be read as "zwanzigeins", as in their name.

According to the association, the inverted pronunciation of the German numbers may have negative consequences in several areas:

- Education: it may create difficulties in the learning of mathematics at primary school level;
- Business: there might be a high error percentage in the communication of data and figures, especially in foreign trade, which may cause financial losses;
- Politics: a reform may be required to adapt to international standards and those used within the European Union; moreover, many foreigners have difficulties in learning German numbers.

3 http://www.verein-zwanzigeins.de/

The association does not, however, aim to abolish the current pronunciation completely, but intends to promote a parallel usage of the two, perhaps following the example of the Czech language, where both varieties coexist.

The problem is now exacerbated by the digital and technological era: typing a (German) multiple-digit number on a keyboard is not an easy task, as it is difficult to leave an empty space to the left to type in the second digit when it is uttered (as many native speakers use to do when writing longhand). This has obvious practical implications, as when hearing a 5-digit number one has to type the second digit first, remembering the first one and typing it afterwards, then type the third one, remembering the fourth one when typing the fifth, and so on.

Hence, one of the slogans of the association is:

> Sorgen wir für Erleichterung und beseitigen wir für unsere Nachkommen alle Schwierigkeiten, die nicht in der Natur der Sache liegen!

3. Simultaneous Interpretation and numerals

SI is a complex cognitive task during which the interpreter has to carry out several operations at the same time or, at least, in very brief succession. Chernov (1994: 140) defines SI as "a complex, bilingual, meaning-oriented communicative verbal activity, performed under time constraints and with a strictly limited amount of information processed at an externally controlled pace". Under such extreme circumstances, not all verbal messages, but only messages with an adequate degree of redundancy, can be interpreted simultaneously. This means that there are several "shadow zones" in SI which are commonly recognized to be particularly difficult even for professional interpreters. The high error score reported by the studies mentioned in the first paragraph suggests that numbers are one such element in SI. This could be due to several universal, language-independent causes which might occur in all language pairs. In the following sections some of the most significant among them will therefore be discussed.

3.1 Universal causes of low accuracy for numbers in SI

3.1.1 The difference in hearing modality

Numerals are managed in a very different way compared to other semantic elements of a phrase. In the listening phase of SI, hearing a numeral causes problems for the interpreter because it has to be perceived both integrally and correctly, contrary to other phrase elements, which can be reconstructed or deduced from the context. Whereas for other elements the interpreter's attention is focused on the semantic meaning, dealing with numerals "intelligent hearing"[4] has to be abandoned in favour of "literal hearing". This breaks the usual mental activity required during SI because, as Seleskovitch (1975: 126) says, "*le chiffre, intervenant à brûle-pourpoint dans le raisonnement, ramène l'attention sur la perception auditive du discourse.*"

After translating the numeral, the interpreter has to shift back to "intelligent hearing", otherwise running the risk of losing important information following the numeral; this continuous switching between the two modalities and the continuous search for a balance between them could be one of the factors leading to frequent errors in the SI of numerals.

3.1.2 The non-application of common SI strategies

The SI task can only be facilitated and mastered once the interpreter has internalized the particular aspects of SI and is able to manage them by means of specific strategies. Gile (1995, 1997) distinguishes between preparation strategies and strategies to be applied during the conference. The latter are also called "coping tactics" and there are three kinds: the coping tactics of *comprehension* (delaying the response, reconstructing the segment with the help of the context, using the booth-mate's help), of *prevention* (changing the Ear-Voice Span, segmenting the text, taking notes, changing the order of the elements in an enumeration etc.) and of *reformulation* (replacing a segment with a super-ordinate term or a more general speech segment, explaining or paraphrasing etc.).

Last but not least, a major strategy applied especially during the SI of a language pair which does not permit a linear transposition is *anticipation.* Anticipation can be defined as the interpretation of a natural piece of spoken text before it is completely finished by the speaker, not only due to intra-textual clues but also due to extra-contextual elements connected with the semantic level of the text. According to the model of *probability prediction,* as elaborated by Chernov (1994), the indispensable premise for the interpretation to take place is *redundancy.* In other words, redundancy allows and at the same time implies the predictability of the message, permitting its anticipation.

The above mentioned strategies – often applied in order to master the SI task – cannot be employed in interpreting numerals, as the verbal numerical system is characterized by some intrinsic features which prevent their successful application.

For instance, numerals are characterized by:

4 This is the listening to the incoming information for a time lapse that is sufficiently long to understand its meaning (Lederer 1982).

- Absence of redundancy – which prevents the application of the anticipation strategy;
- Absence of predictability – which prevents the application of strategies like the reconstruction from the context or again, anticipation;
- Exact content – which requires the numeral to be listened to carefully, thus preventing the application of reformulation strategies;
- Only one meaning – preventing any reformulation;
- Lack of semantic content – requiring literal hearing and thus preventing reconstruction/anticipation.

3.1.3 The role of memory

Numerals are integrated in the linguistic system that they belong to as their meaning strictly depends on their syntactical and lexical formats. This means that the recall of previously stored numbers will be successful only if each single item composing them can be recalled in the correct order as per the original stimulus. The phonological trace of the number is constantly refreshed by so-called subvocal repetition (reproducing the order in which the items have been heard).

Memorizing numbers, especially longer ones, is particularly difficult in the extreme conditions under which SI takes place; the human working memory (WM) has only a limited storage capacity[5] and exceeding its threshold may lead to either a loss of information or wrong recall. Furthermore, the memory span for digits[6] varies according to the language used, since longer numerical expressions in some languages (e.g. Welsh compared to English) require a longer time to be processed, resulting in a reduced digit span (Ellis 1992). A reduced memory span is also usual in "late bilinguals" who have learnt their second language later in life (Brown & Hulme 1992).

During SI the normal reaction time increases and the so-called articulatory suppression, which prevents subvocal repetition, takes place. Limited WM storage capacity is possibly the reason why longer strings of numbers, composed of several elements to be read in different blocks, are the most difficult to retain and thus to translate correctly. In addition, the analysis of large numbers has shown that even on their own they create a considerable load for the WM. As they are composed of several elements and since each element is independent of the others, considerable WM

5 It has been shown that the average storage capacity is about seven unrelated items (Miller 1956).

6 Several experiments on memory involved memorizing digits, especially as a secondary task, to test capacity during concurrent tasks. The *digit span* can be determined with an experiment in which subjects are presented with lists of digits of increasing length and asked to repeat them. When repetition is accurate 50% of the time, the digit span has been reached (Gran 1997).

capacity must be allocated to memorize their correct sequence (i.e. a number like 638.146 needs 13 lexical elements to be transcoded in the German phonological verbal code: *"sechs-hundert-acht-und-drei-ßig-tausend-ein-hundert-sechs-und-vier-zig"*).

The effect of large numbers on WM is therefore likely to be greater than that of small numbers during cognitive tasks and thus to increase the WM effort, being represented by a long string of items that must all be retained in order to interpret the number correctly.

3.2 Particular causes of low accuracy for numbers in SI

The features described in the previous sections and confirmed by previous studies suggest that numbers are problem triggers in SI. The working hypothesis underlining this study is that apart from the above-mentioned language-independent causes of disruption in the SI of numerals, the performance might also be influenced by some language-dependent factors. These are represented essentially by the different linguistic formulation of numerical structures (e.g. the application of the inversion rule in German, but not in Italian or English) compared to Arabic numbers and the length of the German verbal numerals, which requires a longer processing time.

During SI several cognitive tasks have to be carried out at the same time, requiring a great cognitive effort. An interpreter with little experience, like the participants in this study (student interpreters), or even the more experienced, may therefore make the mistake of following the phonological mapping of the SL in their interpretation into the TL. In the case of SI from German into Italian this can lead to inversion errors or, more generally, to transposition or position errors.

4. The experimental study

4.1 Aims

The present study aims to analyse and draw conclusions upon:

- What difficulties, if any, are encountered by a student interpreter in the SI of a speech dense with numbers, like the one used in this study;
- Whether the results for the two language pairs show similar trends, based on a series of statistical analyses;
- Which class of numbers is more prone to produce errors, as the assumption is that different numeral typologies lead to different kinds of errors;
- Which typology of error is the most common;

- Whether student interpreters prefer to note down numbers in the booth or not – if given a choice – and whether note taking reduces the error rate.

4.2 Subjects

The present experiment was conducted on 16 students attending the SSLMIT[7] in Trieste. All participants were Italian mother tongue speakers and had attended at least two years of interpretation courses at the SSLMIT. All subjects were female,[8] aged between 23 and 28 and right-handed.

The 16 students were divided into two groups, one group consisting of eight students having German as their first foreign language (B-language) and the other one consisting of eight students having English as their B-language.

A fundamental criterion for the selection of the participants was a successful result in the SI exam from their B- into their A- language (i.e. German into Italian for the first group, English into Italian for the second one), as the study text was to be interpreted from German into Italian by the first group and from English into Italian by the second group.

4.3 The experimental text

The ideal text for the study had to be dense with numbers but at the same time it had to be presented in a spoken format in order to simulate real working conditions[9] as much as possible.

The text chosen was an actual speech given by the CEO of a well-known German automobile producer at its 2006 Annual Press Account Conference. The speech was originally in German, but the official translation into English was accessible on the company website – and was proofread by a mother tongue editor. Hence, it was possible to use two perfectly equivalent texts, one in German and one in English, both containing exactly the same amount (61) and type of numbers.

Some preparatory texts were distributed to the participants a few days before the experiment, along with a glossary of the most difficult

7 Scuola Superiore di Lingue Moderne per Interpreti e Traduttori.

8 No male subjects were found who satisfied the required selection criterions.

9 It is nevertheless unavoidable that an experimental situation differs from a real life one, as noted in the studies conducted by Alessandrini (1990), Crevatin (1990), Braun and Clarici (1997) and Mazza (2000). The experiment was artificially set up, it did not take place in real working conditions, the participants were students and not professionals, the speech was recorded and not spontaneous.

expressions, in order to allow student interpreters to prepare for the SI – as professional interpreters usually have the chance to prepare themselves for a conference.

The two equivalent study texts were read and recorded by mother tongue speakers in order to give the exact same speech to all subjects of each group. The length of the speech was 8'12" in German and 8'33" in English.

4.4 Methodology

Numbers contained in the texts were classified into five categories according to their size and type:

A. Numbers with 4 or more digits read at once (i.e. 920,000);
B. Numbers with 4 or more digits read in two blocks (i.e. 928,346);
C. Numbers with less than 4 digits;
D. Decimals;
E. Dates.

A categorization of number errors was also set up, partly based on that identified by Braun and Clarici (1997) and partly adapted for this study which, contrary to the previous ones, has a cross-linguistic dimension.

The typologies of number errors identified for this study are:

1) *Omissions*: the numeral is left out altogether or replaced by a generic expression such as *molti, pochi* (many, few), etc.;
2) *Approximations*: although the translation respects the right order of magnitude, it is rounded up or down. The interpreter is usually aware that the SL number was different and accompanies his/her interpretation with a lexical element (e.g. 47,325 being translated as *più di 47.000*, "more than 47,000" or 8.1% being translated as *8% circa*, "about 8%"). However, approximation cannot be considered an error of the same severity as the others, as the message conveyed is not altogether wrong with respect to the stimulus. Several authors do define approximation as a useful strategy or *éscamotage* to overcome translation difficulties, stating that the most important thing is to convey at least the right order of magnitude. For the purposes of this study, it was therefore decided that if a number was affected by two different mistakes (for example approximation *and* lexical error) it would be categorized by the error type which would most change the original communicative intention.
3) *Lexical errors*: the order of magnitude of the stimulus is maintained, but one or more number-words within the numeral have been misinterpreted (e.g. 277,000 translated to 276,000, or 2004 translated to 2005).

4) *Syntactical errors*: the number is of a wrong order of magnitude even if possibly containing the right figures in their correct sequential order (e.g. 300 being translated as 300,000, 150,000 translated to 1,500 or 47,000 to 47%).
5) *Errors of phonemic perception*: the error can be related to a phonemically wrong perception of the *stimulus* in cases of similar sounding linguistic features (e.g. 17, "seventeen", perceived as 70, "seventy").
6) *Errors of transposition (of the digits) or position errors*: the wrong assembly of the figures composing the number, which are correctly selected but misplaced. This includes on the one hand the classical inversion errors typical of the German-Italian language pair with its different numerical structures, but on the other hand it also includes all position errors, possible in English as well, which are not directly attributable to the numerical system structure. The extension of the typical category of inversion errors (identified also by Braun and Clarici in their study, 1997) to a broader error category, which could count for such mistakes made in English too, was chosen because of the comparative character of this study. In order to carry out the statistical analysis on the final data it was necessary to make use of error categories applicable to both languages; in some cases the substitution of a digit could not be classified as a simple lexical or inversion error but rather as a transposition or position error, as shown in the following example:
   7.6% → 6.7%
   8.1% → 1.8%
   528,015 → 285,000.
7) *Other mistakes*: this category includes all other mistakes not belonging to any of the previous types and whose causes are often not apparent. These errors are kept apart and form a rather miscellaneous group (e.g. 528,015 translated with 270,000 or 22.4% with 3.5%).

4.5 Procedures

The experiment took place on two consecutive days at the SSLMIT in Trieste. The two experimental texts had been read and recorded by two mother tongue lectors with a digital double track recorder (DAT) SONY TCD-D7 while a person sitting next to them monitored the speech speed. Every subject was given a piece of paper and a pen and was free to decide whether to take notes while interpreting.

After the SI each subject was asked to hand in the piece of paper if notes had been taken and to fill in a questionnaire about his/her perception of the source text and his/her performance.

At the end of the trials the material was collected and the transcription phase of the 16 interpretations begun, according to the methodology

10 In order to maintain consistency with all experimental studies carried out at the SSLMIT.

established by the *"Dipartimento di Scienze del Linguaggio, dell'Interpretazione e della Traduzione"* of the University of Trieste.[10]

4.6 Statistical analyses

4.6.1 Hypothesis testing

In order to establish how relevant *the language factor* in the SI of a speech dense with numbers is, two mutually exclusive hypotheses were formulated. The first hypothesis – H0 – argues that the language factor is relevant in the SI of a speech dense with numbers and the second one – HA – argues that the difference is not significant.

To verify which of the two hypotheses should be rejected, the unpaired Student's t-test for independent samples (C.I. 95%) was carried out, being one of the most commonly used techniques for testing a hypothesis on the basis of a difference between sample means.

In this study it was used to:

- verify which hypothesis between H0 and HA can be rejected;
- analyse if there are significant differences between the results obtained for the German Text (GT) and for the English Text (ET) concerning the general proportion of errors in each numeral category;
- analyse if there are significant differences among the results obtained for GT and ET concerning the trend of the different typologies of mistakes made.

A further statistical parameter – correlation – was used in order to analyse the positive or negative correspondence between note taking and the performance.

5. Results

5.1 General performance in GT and ET

At the end of the trials, all wrongly interpreted numerals were counted to assess the subjects' performance.

From an initial analysis, it was striking that not one of the study subjects interpreted 100% of the numerals correctly. The error score on the total amount of numbers in the texts corresponded to 40.6% (mean value: 24.8) in the GT and 41.2% (mean value: 25.1) in the ET. These figures were considerably high and in order to verify if their difference was statistically significant, the Student's t-test was conducted on the mean values, with the following outcome:

[Stat t = -0.156; df = 14; p = 0.87].

The resulting p> 0.05 indicates that the difference is *not* significant. From this first overview, the H0 hypothesis, according to which the language factor is relevant in the SI of a speech dense with numbers, can be rejected in favor of HA.

However, the comparison between the general performances in the two texts indicates only a general trend. For this study, different numeral categories and different error typologies have been taken into account as well as whether a detailed analysis of these factors could lead to a partial re-evaluation of the first results obtained.

5.2 Errors and numeral classes

As pointed out previously, the error score was 40.6% for the GT and 41.2% for the ET. These total error scores affected the five numeral categories to varying extents. For each category, the mean score of the errors was calculated to determine which numeral category was most prone to errors. The results are illustrated in figure 3 below.

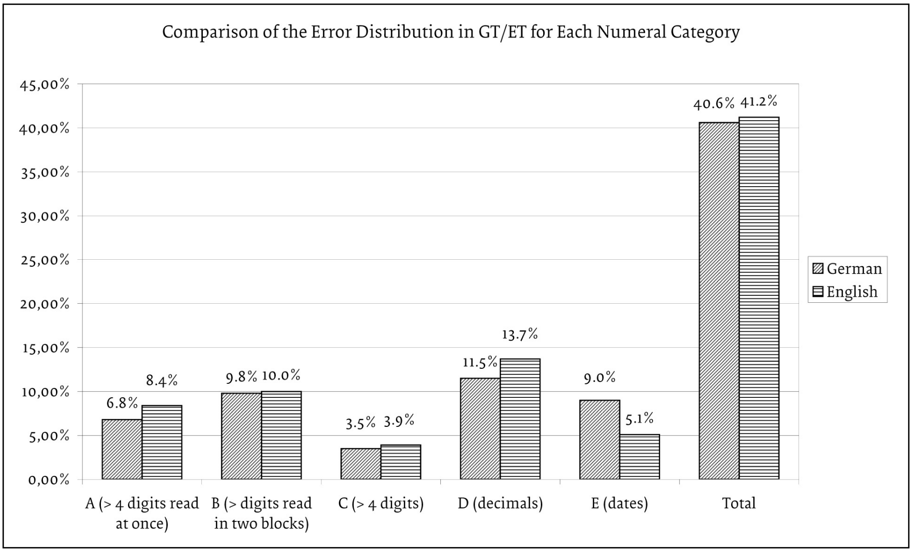

Figure 3.

From an initial analysis, most mistakes involved the numeral classes D (decimals) (11.5% of the total mistakes made in GT and 13.7% in ET) and B (numbers with 4 or more digits read in two blocks) (9.8% in GT and 10% in ET), while the most accurately interpreted class turned out to be C (numbers with less than 4 digits) (3.5% and 3.9% in GT and ET respectively). Category A (numbers with 4 or more digits read at once) showed a higher error score in ET than in GT (8.4% vs. 6.8%) thereby becoming the third most difficult category to interpret after D and B, showing a slightly higher, though not significant, difference between the two languages.

This parallelism was only interrupted by the results obtained for category E (dates) where errors made in GT exceeded those made in ET. The dates of the study text often referred to the current year of business and were therefore repetitive and redundant. This feature required a different processing modality, which might have led to some different strategic choices and hence results.

The above described trend was confirmed by the Student's t-test:

Category	Stat t	df	P		Outcome
A (whole ≥ 4 digit 1 block)	-0.984	14	0.34	(p > 0.05)	Difference is not significant
B (whole ≥ 4 digits 2 blocks)	-0,260	12	0.79	(p > 0.05)	Difference is not significant
C (< 4 digits)	-0.361	14	0.72	(p > 0.05)	Difference is not significant
D (decimals)	-0.955	14	0.35	(p > 0.05)	Difference is not significant
E (dates)	2.850	14	0.01	(p < 0.05)	*Difference is significant*

While figure 3 shows the major differences or similarities given the same conditions, representing the objective distribution of the total errors made among the numeral categories, figure 4 examines the error score in relation to the number of items in each category, showing clearly which categories are more susceptible to mistakes:

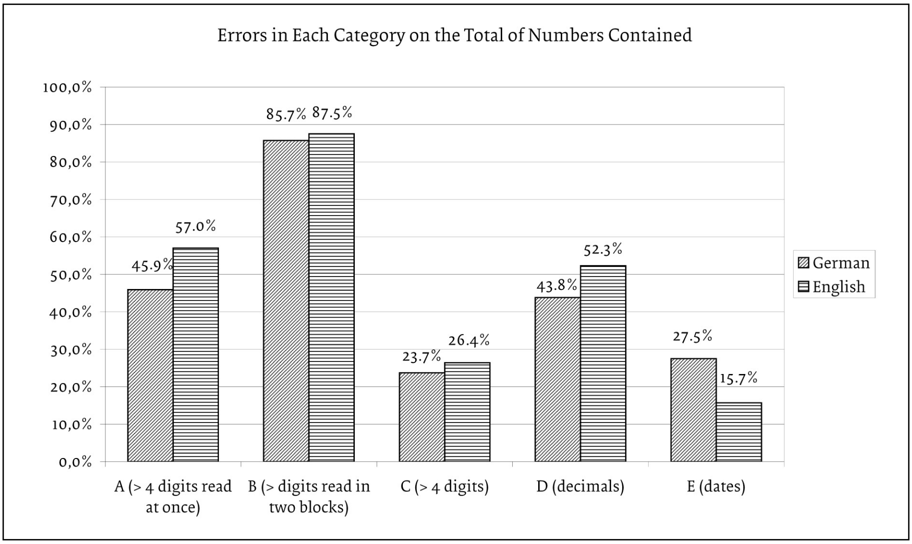

Figure 4.

Class B (numbers with 4 or more digits read in two blocks) proved to be the least accurately interpreted category (85.7% in GT, 87.5% in ET). Category A (numbers with 4 or more digits read at once) was misinterpreted in approximately half the cases (57% in ET and 45.9% in GT), followed by D (decimals, 52.3% in ET and 43.8% in GT); C (numerals with less than 4 digits) and E (dates) were interpreted more accurately.

5.3 Error typologies

The total error scores (40.6% in GT and 41.2% in ET) were made up of different types of errors; figure 5 summarizes their occurrence in the two texts:

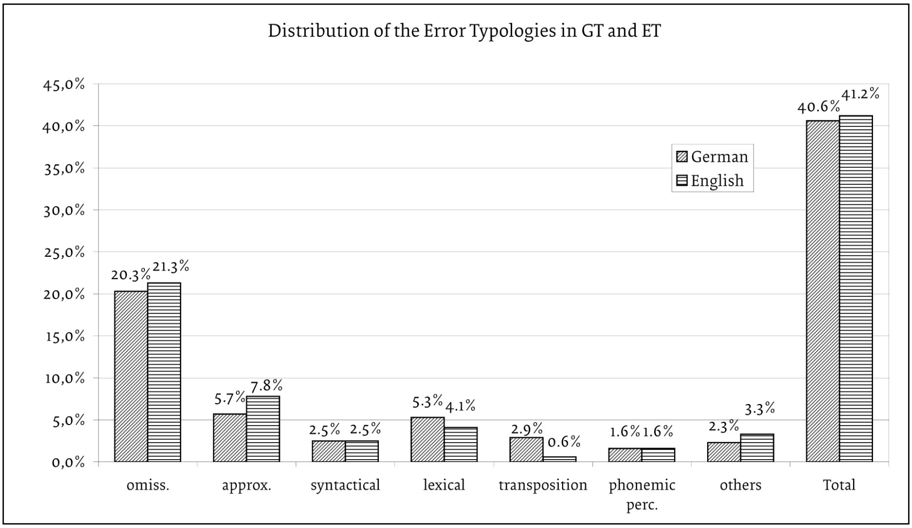

Figure 5.

Some error typologies – particularly omissions – showed a very high error percentage (20.3% in GT and 21.3% in ET of the total amount of errors made). The other error typologies were distributed quite uniformly: they ranged from 7.8% (approximations in ET) to 0.6% (errors of transposition, again in ET). This suggests that the numerals contained in the study texts were affected by different error typologies and that the underlying causes were different.

To allow a clearer and direct comparison between GT and ET, the breakdown of the different error typologies was calculated on the total number of errors made, as shown in figures 6 and 7 (breakdown per language) and in figure 8 (comparison between German and English):

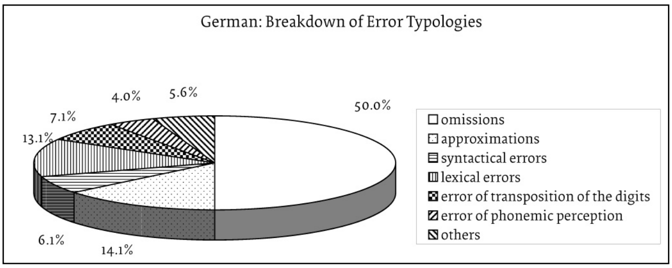

Figure 6.

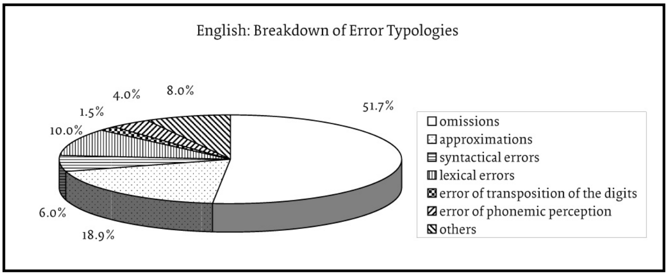

Figure 7.

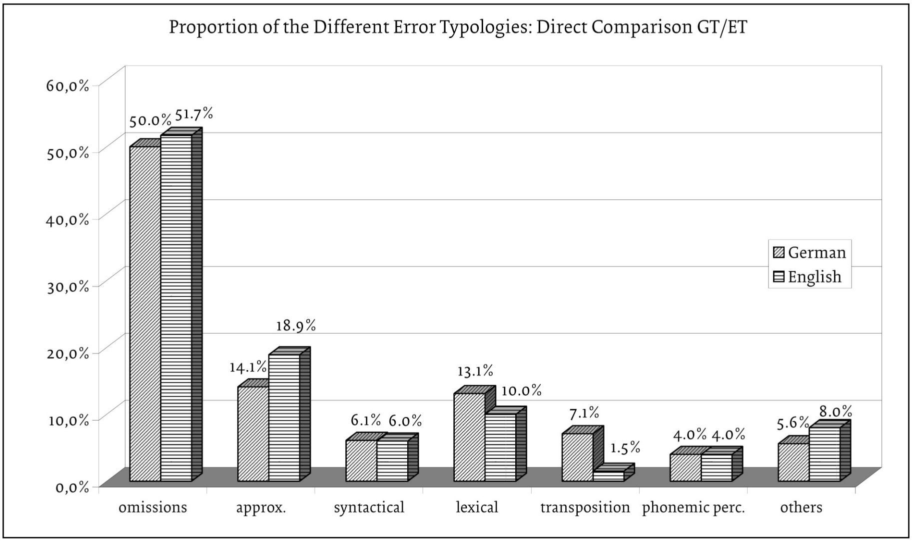

Figure 8.

Omission was the most common mistake in both texts (50% in GT and 51.7% in ET of the total errors made). Approximations were also fairly frequent, but slightly higher in ET (18.9% vs. 14.1%), whereas the transposition errors were rather higher in GT (7.1% vs. 1.5% in ET). Lexical and syntactical errors, those of phonemic perception and other mistakes followed a homogeneous trend, with a slight prevalence of lexical errors in GT (13.1% vs. 10%) and of other errors in ET (8% vs. 5.6%).

These results were verified through the Student's t-test:

Error Typology	**Stat t**	**Df**	**P**		**Outcome**
Omission	-0.290	13	0.77	p> 0.05	Difference is not significant
Approximation	-1.091	13	0.29	p> 0.05	Difference is not significant
Syntactical	0	14	1	p> 0.05	Difference is not significant
Lexical	1.287	14	0.21	p> 0.05	Difference is not significant
Transposition	2.33	8	0.04	p< 0.05	*Difference is significant*
Phonemic Perc.	0	7	1	p> 0.05	Difference is not significant
Others	-0.828	14	0.42	p> 0.05	Difference is not significant

The results obtained through the Student's t-test confirmed that the only significant difference between the two texts as far as error typologies were concerned was represented by the errors of transposition (or position errors). These included all the cases in which the error consisted of a wrong assembly of digits within the number, including all inversion errors in the classical sense – i.e. the inversion of teens and units. The proportion of inversion errors in the transposition category led to the higher incidence of transposition errors in German compared to English. For instance, in the latter language such errors were determined not by internal factors, but by external ones, such as a sort of echo effect on the numeral itself.

6. Note taking

In this study the participants had the choice of taking notes during the SI if they so wished. The only condition was that they had to hand in the papers if they wrote down notes, so that they would be analysed in order to establish whether note taking was a valid support in the SI of numbers or if the inaccuracy in numbers could be attributed to that very action.

Only one participant chose not to take any notes (SI from German into Italian).

The first analysis consisted of counting how many numbers were noted down, how many of them were written down correctly, and the correspondence between note taking and performance.

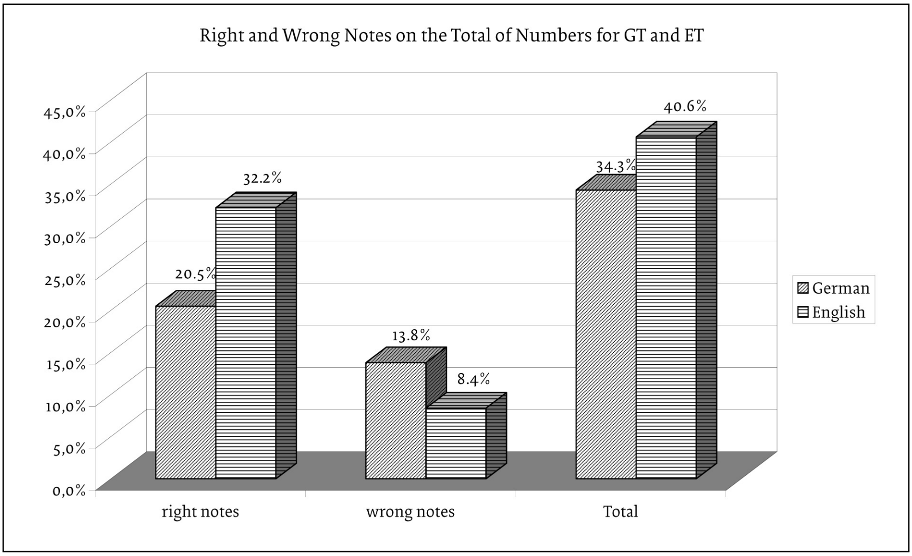

Figure 9.

Figure 9 shows that more notes were taken in the SI from German into Italian than from English to Italian (40.6% against 34.3%). This could be ascribed to the awareness of the interpreters from German regarding the different structures of the numerical systems. The GT group, conscious of the intrinsic inversion difficulty, made recourse more often to paper support than the ET group by means of Arabic notation. The mental application of the inversion rule would be more energy consuming, especially for subjects who did not learn German as their mother tongue. This difficulty, already apparent in a normal communicative situation, becomes more crucial in SI, which takes place under time constraints not normally present. The widespread use of note taking is understandable: writing the numbers down can help the interpreter feel more confident and able to avoid the mistake of following the phonological mapping of the stimulus, as first the units are written down and then the teens.[11] Afterwards, it can be read out more effortlessly at the moment of reproducing the numeral in the target language.

German notes were more prevalent and generally more correct than English notes. This is possibly due to the fact that interpreters working from German are more used to taking notes, as a more or less automatic operation, being aware of the internal structural difficulty. This was not always the case in English, where the difficulty may be perceived as something "external", thus leading to poorer application of prevention tactics (and to less practice in them, determining more errors). Only two participants in the group interpreting from English into Italian had German in their linguistic combination, which may be relevant.

The breakdown of the wrong and right notes taken is summarized as below (figures 10 and 11):

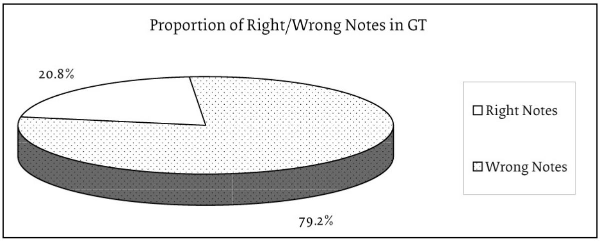

Figure 10

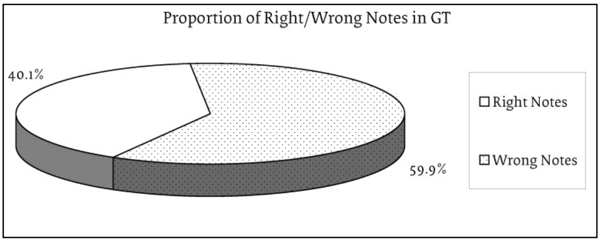

Figure 11

11 As stated by almost all the students in the questionnaire given at the end of the trial.

As so many notes were taken, it was felt necessary to analyse whether they turned out to be a valid support to the SI of a speech dense with numbers.

Of the total notes taken in GT, 71.2% were right leading to correct interpretation, 8.1% right leading to wrong interpretation, 19.7% wrong leading to wrong interpretation and 1% wrong notes leading to correct interpretation.

In ET the proportion was slightly different: of the total notes taken, 56.9% were right leading to correct interpretation, 3% right leading to wrong interpretation, 39.5% wrong leading to wrong interpretation and 0.6% wrong leading to a right interpretation.

There was a general linearity among the results; if a number was correctly noted, it was often interpreted correctly and vice-versa. This was also confirmed by the analysis of the correlation between notes and performance:

The coefficient *r* of correlation for the two texts was:

GT	+0.76
ET	+0.89

The two values were close to +1: this means that the variables "notes and performance" were positively correlated, so when one increases, the other one does the same. The slightly lower value obtained in GT is due to the fact that in 8.1% of cases right notes led to a wrong interpretation.

Subjects tended not to manage complex numbers well, such as those belonging to category B (numbers with 4 or more digits read in two blocks), which were presumably perceived as structurally too complicated even to write down (only 31% of the total amount of numbers present in that category in GT and 33.4% in ET were noted down). A and D were the categories most frequently noted down (A: 57% in GT and 63.8% in ET and D: 60.1% in GT and 51.6% in ET) as they are quite complex to retain in memory but quite easy to represent with the Arabic code (requiring just a few digits). C and E were the categories least noted down: C (less than 4 digits), as they are very short and thus easier to remember (47.2% of the total amount in GT and 29.1% in ET) and E (dates) as they are perhaps the simplest to interpret, requiring a different processing modality (13.1% in GT and just 4.3% in ET). The dates category is presumably the only one that can be "visualized" and associated to a semantic meaning, which makes retention simpler than in the others, where interpreters have to rely only on phonological clues.

7. Discussion

The results suggest a fairly homogeneous trend for the two language pairs analysed, but at the same time also show some differences. The most obvious similarities are the results concerning the omissions, which turned out to be the most common mistake in both texts, followed by approximations, and the results concerning category B (numbers with 4 or more digits read in two blocks), which got the highest error score. As expected, the major difference occurred in the error category "transposition of the digits" or "position errors", which were much more prevalent in German than in English.

In terms of the high error score for omissions, reference can be made to attention factors. When a numeral is inserted into a text it is necessary to adopt a different strategy in order to comprehend and translate it successfully. For this reason, two different kinds of memory have to be concurrently activated: the semantic one for the text and the literal one for the numerals. The effort required to accomplish this operation can lead to an overflow and thereby loss of information (omission). There are two types of omissions: deliberate omission, i.e. when the interpreter decides deliberately to omit a problematic part of a textual segment, and omissions due to the exhaustion of the cognitive resources allocated for the listening and analysis phase.

Finally, this result is in line with the previous studies, where omissions also proved to be the most common error.

The category which showed the highest error score in relation to the number of items it included was category B (numbers with 4 or more digits read in two blocks), which could well be due to the word-length-effect on WM. Longer words are more difficult to retain after a brief exposure to information and the memory span decreases when the words to recall are long. Moreover, sub-vocal repetition, intended to refresh the mnemonic trace, does not take place in SI because of articulatory suppression. This could account for the high error score found in numbers belonging to category B, the longest and the most difficult to retain. Deloche and Seron (1982: 125) also proved that errors increase according to the length of the numeral, stating that *"the error rate increases continuously as a function of the number of words in the numeral. A purely quantitative analysis thus indicates that the length of the numeral is a pertinent difficulty factor."*

Concerning the word-length-effect on SI, Ellis (1992) stated that the time necessary to utter a numerical expression influences the ability, even of a native speaker, correctly to recall the number and that the memory span thus further decreases when numbers are expressed in a language whose verbal codification requires more time. The results of this study, however, diverge slightly from this conclusion. A number expressed in the German verbal phonological code requires more time to be articulated than the same expression uttered in English, as it is composed of more or longer syllables. However, in this study, large numbers were affected more or less by the same error score in both languages. The word-length-effect on SI was therefore significant when dealing with B category (numbers with 4 or more digits read in two blocks) compared to C (numbers with less than 4 digits), but not from a cross-linguistic point of view.

The prevalence of transposition errors in German can primarily be ascribed to the incidence of classical inversion errors in this category. It can be assumed that the inversion of numerical expressions causes a higher degree of difficulty in their transcoding as far as non-native speakers are concerned. For instance, it is harder to acquire the habit of codifying a number system that does not follow the decreasing ordering of magnitude present in the numbers (...tens, teens, units – which is the case in German) on the basis of a linear system already acquired (which is the case in Italian mother tongue speakers), as this means changing from linearity to non-linearity.

8. Final remarks

8.1 About the methodology: innovations of the present study

The present experiment was especially set up to investigate the differences and similarities of a SI in the context of a speech dense with numbers from German into Italian and of the same speech from English into Italian.

The methodology adopted showed some innovations compared to previous studies as a *cross-linguistic* analysis with categorization of numbers and error typologies was previously untried.

The classification of the number categories used stemmed largely from an empirical observation: if it was true that big numbers caused more problems in SI, it was necessary to further specify what "big numbers" meant: for instance, there are numbers that need five digits to be represented in the Arabic code, but contain many more lexical elements to be transcoded into the phonological code, and other numbers which also require five digits in the Arabic code but fewer lexical elements (i.e. 39,000 and 39,754: they are both represented by five digits in the Arabic code but the first one requires five lexical elements to be transcribed in the German verbal phonological code – "neun-und-drei-ßig-tausend" – and four lexical elements in the English one – "thir-ty-nine-thousand" –, whereas the second number requires eleven lexical elements in the German verbal code – "neun-und-drei-ßig-tausend-sieben-hundert-vier-und-fünf-zig" – and ten in the English one – "thir-ty-nine-thousand-seven-hundred-and-fif-ty-five"). This aspect could not be neglected when setting up the number categorization which led to the distinction made between categories A (4 or more digits read at once) and B (4 or more digits read in two blocks) – with reference to the pause in correspondence to the multiplicator "thousand", which for the first numeral indicates its end and for the second one indicates the beginning of a new functional unit.

Concerning the classification of error typologies, a special category was set up for this study: that of "transposition of the digits", or "position errors". This stemmed from the cross-linguistic dimension of the present study and from the observation of errors which could not be categorized either as lexical errors or classical inversion errors.

Finally, the study participants could freely choose whether or not to take notes when interpreting simultaneously. However, the general degree of inaccuracy for numbers is rather similar to other studies where this was partially forbidden in some trials, suggesting that there are no specific strategies that can ensure a definitive solution of the problem.

8.2 Conclusions

Data collected from the experiment suggest that inaccuracy for numbers by student interpreters was rather high both in the SI from German into Italian (40.6%) and in the SI from English into Italian (41.2%). This confirms the theory that numbers are disruptive elements in SI. The Student's t-test conducted on these values stated that the difference was not significant and that the difficulty is *language-independent*. The hypothesis according to which the *language factor* is relevant can so far be rejected.

However, starting from similar conditions, the breakdown of the error typologies did represent a main difference in the results and this led to a partial re-evaluation of the outcomes.

The significance of the Student's t-test as far as the errors of transposition of the digits in German were concerned represented the most significant aspect. This result was counterbalanced by the preponderance of other types of errors in English, most of all by approximations, which seemed to be due to external causes. As pointed out previously, however, approximation errors in English did not represent a severe mistake, at least not from a semantic point of view – unlike transposition errors, which were much less significant for the English-Italian language pair.

Several patterns were accounted for by external factors: factors linked to the particular textual nature, clusters of numerals in certain parts of the text and so on. In German, apart from these aspects, there was an additional hurdle represented by the intrinsic internal difficulty posed by the inverted numerical system, which required even greater concentration.

Given all the conditions described in the previous paragraphs, the study has hopefully contributed to shedding light on this particular aspect of SI, which is still under debate and rather controversial.

The degree of accuracy for numerals was quite low; there does not seem to be any real solution to this problem, apart from practical expedients such as note taking and the help of the booth mate. Perhaps the problem should be solved at the origin and, as Pearl suggests:

> [...] speakers would be well advised before using figures to reflect on whether their point could just as well be made by giving an order of magnitude, such as: 'much', 'little', 'few', 'a tremendous amount', 'sufficient' etc. (1999: 21).

In addition, specific training in this kind of text typology could help coping with the problem: as Baddeley (1990) says, "practice makes perfect!"

References

Alessandrini M.S. (1990) "Translating numbers in consecutive interpretation: an experimental study", in *The Interpreters' Newsletter*, 3, 77-80.

Baddeley A.D. (1990) *Human Memory: theory and practice*, London, Lawrence Erlbaum Associates.

Bosshardt H.G. (2004) *Psychologische Vorgänge bei der Repräsentation von Zahlen* (Manuskript des Vortrags gehalten auf dem Interdisziplinären Kolloquium "Das deutsche Zahlenaussprechsystem", 19th January 2004, Bochum).

Braun S., Clarici A. (1996) "Inaccuracy for numerals in simultaneous interpretation: neurolinguistic and neuropsychological perspectives", in *The Interpreters' Newsletter*, 7, 85-102.

Brown G.D.A., Hulme C. (1992) "Cognitive psychology and second language processing: the role of short-term memory", in R.J. Harris (ed.), *Cognitive Processing in Bilinguals*, Amsterdam, Elsevier, 105-121.

Chernov G.V. (1994) "Message redundancy and message anticipation in simultaneous interpretation", in B. Moser-Mercer and S. Lambert (eds.), *Bridging the Gap: Empirical Research in Simultaneous Interpretation*, Amsterdam/Philadelphia, John Benjamins, 139-151.

Crevatin A. (1991) *La traduzione dei numeri in interpretazione simultanea: un contributo sperimentale,* unpublished thesis, SSLMIT, Università degli Studi di Trieste.

Dehaene S. (1992) "Varieties of numerical abilities", in *Cognition*, 44, 1-42.

Deloche G., Seron X. (1982) "From one to 1: an analysis of a transcoding process by means of neuropsychological data", in *Cognition*, 12, 119-149.

Deloche G., Seron X. (1987) "Numerical transcoding: a general production model", in Lawrence Erlbaum Associates (ed.), *Mathematical Disabilities. A Cognitive Neuropsychological Perspective,* New Jersey, Hillsdale, 137-170.

Ellis N. (1992) “Linguistic relativity revisited: the bilingual word length effect in working memory during counting, remembering numbers, and mental calculation”, in R.J. Harris (ed.), *Cognitive Processing in Bilinguals*, Amsterdam, Elsevier Science Publishers B.V., 137-155.

Gile D. (1995) *Basic Concepts and Models for Interpreter and Translator Training,* Amsterdam/Philadelphia, John Benjamins.

Gile D. (1997/2002) “Conference interpreting as a cognitive management problem”, in F. Pöchhacker, M. Shlesinger (eds.) (2002), *The Interpreting Studies Reader,* London/New York, Routledge Language Readers, 163-176.

Gran L. (1997) “Introduzione ai meccanismi della memoria”, in L. Gran and A. Riccardi (eds.), *Nuovi orientamenti negli studi sull’interpretazione,* Sert 6, Trieste, SSLMIT, 1-18.

Hurford J.R. (1987) *Language and Number. The Emergence of a Cognitive system*, Oxford, Basil Blackwell.

Lederer M. (1982) *La traduction simultanée, experience et théorie,* Parigi, Minard Lettres modernes.

Mazza C. (2001) “Numbers in Simultaneous Interpretation”, in *The Interpreters’ Newsletter,* 11, 87-104.

Miller G. (1956) “The magical number seven, plus or minus two: some limits in our capacity for processing information”, in *The Psychological Review,* 63, 81-97.

Pearl S. (1999) “The other three eights and the four 'F's: finiteness, fallibility, freedom of speech and fair competition in the simultaneous interpretation environment”, in *The Interpreters’ Newsletter,* 9, 3-28.

Seleskovitch D. (1975) *Langage, langue et mémoire. Etude de la prise de notes en interpretation consécutive,* Paris, Minard Lettres Modernes.