Transcribing speech:
Errors in corpora and experimental settings

Dr Isabella Chiari
Dipartimento di Studi filologici, linguistici e letterari
University of Rome "La Sapienza"
isabella.chiari@uinroma1.it

1. Introduction

Administrations, government organs, judiciary courts always faced the problem of defining limits in transcription practices. Nowadays corpus linguistics and computational linguistics have focused their attention on spoken corpora as indispensable tools for descriptive linguistics, as well as for applied purposes (in speech technologies, such as text-to-speech and speech recognition, in dialogue systems, in natural language processing and information retrieval, etc.). But transcription is not just a meta-linguistic practice serving linguistic analysis, it is, at the same time, a linguistic act itself, governed by its own strategies and tightly linked to other speech acts, and linguistic practices (such as note-taking, listening to spoken language for different purposes, writing following dictation, etc.). Recent literature has often been centred on transcription system design, on reviewing and comparing different transcription systems (Chafe, 1995; Connell & Kowal, 1999; Cook, 1995; Derville, 1997; Edwards & Lampert, 1993; Lapadat, 2000; Leech, Myers & Thomas, 1995; Pallaud, 2003; Romero, O'Connell & Kowal, 2002), and on errors and inconsistencies in linguistic annotation. Furthermore a large tradition in transcription is common in ethnographic studies (Powers, 2005; Vigouroux, 2007) and in conversation analysis (Ashmore & Reed, 2000).

Transcription of speech is often driven by different transcribers' understanding strategies, leading to specific error typologies (Chiari, 2006a; Chiari, 2006b; Lindsay & O'Connell, 1995; Pallaud, 2002; Pallaud, 2003). How does the transcriber contribute to the reconstruction or mis-reproduction of the spoken text? An analysis of common errors derived from experimentally induced transcriptions and from spoken reference corpora of the Italian language are compared and analyzed quantitatively and qualitatively in order to observe different patterns, relative frequencies, and motivations of occurrence.

2. The practices of transcription of speech

While errors elicited in a controlled experimental environment let us observe transcribers' behaviour horizontally (testing different listening conditions, monitoring the number of repetitions, utterance length, transcribers' sociolinguistic features), corpus transcription evidence provides information on different strategies in a more naturalistic and autonomous context (with indefinite utterance length listening, indefinite number of repetitions, different settings and generally no information on the number and characteristics of transcribers and revisers). Transcription errors are further analyzed in order to find out cognitive and linguistic features that they might share with slips of the tongue, slips of the ear and different typologies of linguistic errors (and ordinary linguistic practices).

Transcription errors collected with different methodologies can constitute interesting evidence for the processes of language production, listening and understanding strategies as well as relevant insights for the elaboration of transcription guidelines to be used for the compilation of spoken corpora.

But what kind of errors are the subject of this investigation? Errors in transcription can cover different aspects of the linguistic and meta-linguistic abilities possessed by native speakers of a language: from perception errors, to annotation errors, to the complex process of labelling non linguistic features of conversation and annotating linguistic properties. The focus of this research will be merely on the identification of the words spoken. Roberts Powers (2005: 1) notes that the instruction "transcribe every word" is often ambiguous since it involves a large number of decisions, transcription being a selective process involving a filtering performed by the trascriber (Ochs, 1979). As it will be showed in the following paragraphs, and as any corpus linguist having dealt with spoken transcripts knows by experience, there are a large number of errors that affect the sheer identification of the words spoken. Those errors are hardly detectable, since they generally result in perfectly grammatical and meaningful transcripts of utterances, and are generally *not* due to bad or noisy recordings.

What is the 'error'? What counts as an error? Many who have dealt with error analysis in different fields of linguistics are extremely careful in using the word *error*. Some prefer to use the word *change* (Lindsay & O'Connell, 1995:102), deviation, "breaches of the code" (Corder, 1973: 259). The case of transcripts is a quite complex one, since there actually is always an interpretation lying under every transcription task. When having to evaluate and judge on performance errors, it is often multifaceted and intricate to identify a norm, so the notion of error appears theoretically weak. In the case of transcription errors we tend to have a final inter-subjective agreement among transcribers and revisers at the end of the listening process, or of the repetitions of the recorded portion. Thus we will use this agreement as the norm to define the transcription error.

The investigation on transcription errors is articulated in two sections corresponding to two different ways of gathering information and data about listening and transcribing strategies. The fist section is conducted in an experimental setting, using audio administered by the experimenter in a controlled setting (Chiari, 2006a; Chiari, 2006b), the second section is a corpus-based research on errors appearing in the final transcription of a corpus of Italian spoken language.

3. Errors in experimental settings

Each participant was submitted to the hearing of 22 different utterances to transcribe. Speech from two different typologies was selected to be included in each test. Type A includes accurately read or spontaneous but controlled speech, selected from television broadcast news or public formal speeches, where one speaker was involved and produced the whole utterance, generally at a quite high word per second rate (characteristic of news reading). An example utterance from this typology is: *L'Italia nella morsa del freddo. Temperature in picchiata da nord a sud, miglioramento previsto da mercoledì* (R26: 5.52 secs). Type B includes spontaneous speech, and

conversation turn recorded in various ordinary situations, mainly from real-tv shows. An example utterance from this type is: *Quando ieri è stata fatta la spesa e si poteva fare qualche altra cosa* (R1018: 2.59 secs).

All digital recordings where acquired directly from tv source in February 2006, and segmented into turns (utterance turns or dialogue turns), and saved in wave format to be heard on a compact player or from pc speakers. The selected recording presented the highest quality of audio sound with least background noise possible and no superimpositions to avoid noise interference in the hearing, understanding and transcribing tasks. Each turn contains only one speaker's voice, and is a full utterance, a brief sequence of utterances or a meaningful portion of a long utterance. Length varies from around 1.5 sec to 13 seconds. Utterances selected for each test typology were chosen to be belonging to the same "spoken text" where possible, as to preserve the listener's ability to rely on what has been previously heard, letting the transcriber find the least artificial condition as possible.

Before each of the two series of hearing exposures, participants were presented with a test for volume adjustment with utterances not belonging to their test type. Before the first series two utterances were added (without telling participants) as a training, and were not computed in the results. Each test consisted of 22 different utterances: the first two were the training utterances, followed by 10 utterance form controlled speech and 10 utterances from spontaneous speech (single dialogue turns with only one speaker talking). 100 utterances were tested (50 in type A speech and 50 in type B).

Participants were given a brief sociolinguistic questionnaire and paper for drafts and were asked to transcribe in handwriting the spoken sequences they heard (choosing their own jotting strategies: online or offline), and then to copy their drafts in an ordered form at the end of data exposure. They were also told to write down only the words spoken (excluding vocal activities, noises and pauses) and not to clean up text, in particular signalling repetitions they heard and not correcting errors produced by speakers. After the data exposure phase participants were not allowed to correct their first draft.

The administration of spoken data was conducted by the experimenter with the aid of a computer with speakers. Before each utterance, participants were told how many times they were to ear it (one to three times depending of length of sequence). The entire duration of the experiment lasted about 30 minutes for each participant. The listening material for each test consisted of about one minute of spontaneous speech and one minute of controlled speech.

Sample spoken material consisted of 100 different utterances (50 in controlled speech and 50 in spontaneous speech), plus two control utterances added at the beginning of the test. The total amount of utterance token presented to the subjects was 400. Utterances ranging one to five seconds were presented once, from five to eight seconds twice, and those lasting more than eight seconds were run three times. Different tests were presented to 20 participants (12 women and 8 men), whose age ranged from 18 to 62 years old, with an average of 28, all having obtained at least an high school degree.

An error analysis was conducted on the transcribed material in order to obtain a full list of errors (*cf.*
Table 1). The 20 utterances belonging to each test were analyzed in order to obtain a full list of errors, where the participant's transcription differed from a supervised transcription (always checked with audio). Missing words or misperception of the first word and last word of each utterance has not been computed, since they involve a certain amount of surprise and voice lowering. Given that participants were not themselves managing repetition of utterances it would have been misleading. A total amount of 455 errors have been reported, with an average of 22.7 errors per participant (about 1.13 errors per utterance heard). 5.75 errors per utterance type were reported in the whole experiment.

Utterances analyzed	400
Errors reported	455
Avrg. nr. errors per participant	22.7
Avrg. nr. errors per utterance	1.13

Table 1 - Results summary

A comparison of different textual typologies was conducted in order to find out if there are any differences in error rate in controlled versus spontaneous speech. Data does not provide any special insight. A slight variation in frequency differentiates the two text typologies selected. Controlled speech induces errors in 48.4% of the total, while spontaneous speech covers 51.6% (*cf.* Table 2). In this specific case since utterances in controlled speech were selected from television news and speeches there is probably an error effect due to fast speech rate of news broadcast reading habits. Usually spontaneous utterances were relatively shorter in duration, and still gathered more errors.

	Frequency	*%*
Controlled speech	220	48.4
Spontaneous speech	235	51.6
Total	*455*	*100.0*

Table 2 - Errors per speech typology

Looking at all the different phenomena together we observe a general tendency at preserving the overall meaning of the sentence (45.9%), especially when single words are affected (and not whole constituents) (55.1% preservations, and 20.7% partial preservations).

	Frequency	**%**
yes	209	45.9
partial	76	16.7
no	170	37.4
Total	*455*	*100.0*

Table 3 - Meaning preservation

Errors were further analyzed to observe more specifically what kind of change occurred in transcriptions. Simple structural categories common in slips and error research were used: substitution, addition, deletion, movement. The most common type of errors were substitutions (205 cases, 45.1%) and deletions (199, 43.7%), while cases of addition (40, 8.8%) and movement (11, 2.4%) were fairly rare (*cf.* Figure 1). A closer observation of change types lets us order and elucidate certain error typologies.

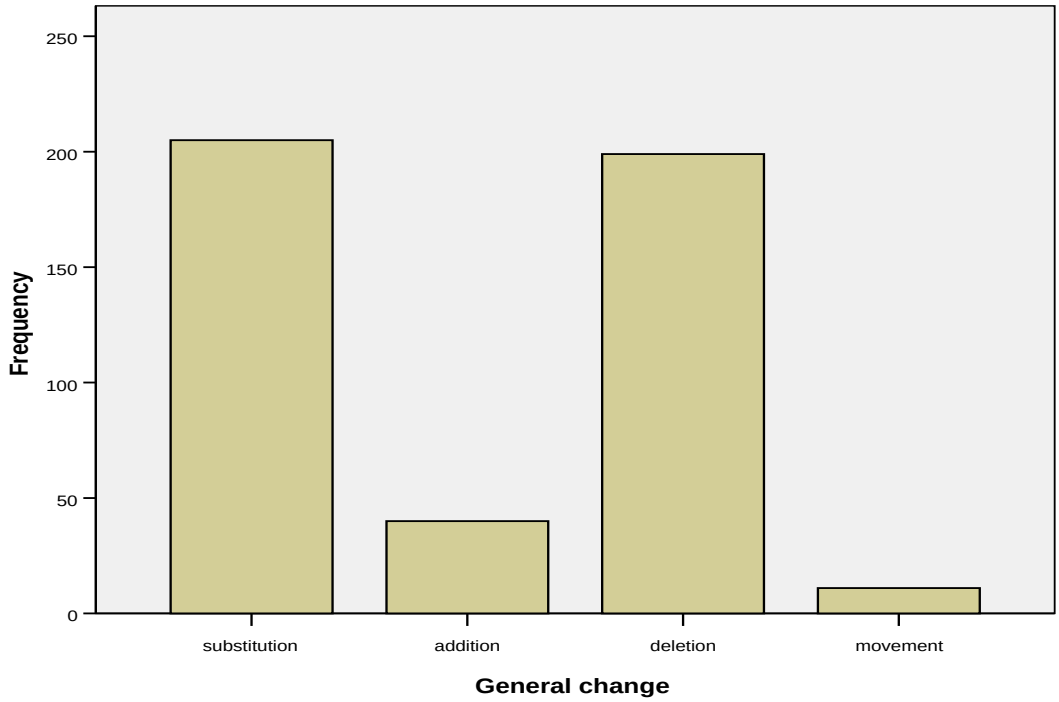

Figure 1 - Type of change in experimental data

Substitutions were an element is switched with another at any linguistic level occurred 205 times (45.1%). Examples are utterances where *un profondo cambiamento* ("a deep change" is transcribed as > *un grande cambiamento* "a great change"; *parla ad un convegno di* ("convention") > *parla a un congresso di* ("congress"); *rendere flessibile il patto* ("make an agreement flexible") > *rendere possibile il patto* ("make an agreement possible")*; scegliere apparecchi di classe A* ("choose appliances of A class") > *scegliere elettrodomestici di classe A* ("choose of household-electric A class")*; che deve fare > che doveva fare*. Among substitutions 52.7% of occurrences involve lexical elements, 19% function words and 16.6% verb conjugation errors (*cf.* Table 4).

	Frequency	*Percent*
lexical switch	108	52.7
sing/plur switch	6	2.9
switch substituent with lexical element	14	6.8
function word substitution	39	19.0
syntactic misplacement	2	1.0

verb conjugation error	34	16.6
phonetic variant	2	1.0
Total	*205*	*100.0*

Table 4 - Substitution typologies

Target grammatical categories (when single words are involved, 137 cases, 67%): verbs 32.1% (44), prepositions 19.0% (27), pronouns 16.8% (23), nouns 14.6% (20), adjective 8.0% (11), adverbs 5.1% (7), conjunctions 2.2% (3), article 1.5% (2) (e.g. Table 5). Substitutions in the great majority of cases involve elements belonging to the same grammatical category (82%). Regarding content preservation in word level substitutions, in 38.7% of cases meaning is preserved completely, in 22.6% is partially preserved, while in 38.7% a complete misunderstanding occurs.

	Frequency	*Percent*
noun	20	14.6
verb	44	32.1
adverb	7	5.1
adjective	11	8.0
conjunction	3	2.2
pronoun	23	16.8
article	2	1.5
preposition	26	19.0
numeral	1	.7
Total	*137*	*100.0*

Table 5 - Target grammatical category of lexical substitutions

Addition or insertion of words is relatively rare (8.8%), and can be generally seen as a repair device where subjects try to give a written textual form to the spoken material (adding conjunctions for examples instead of reporting direct coordination in a sequence of sentences). Examples of additions are: *sentiamo l'inviato > sentiamo ora* ("now") *l'inviato; la verità sapete > la verità lo sapete; ho parlato le ho chiesto > ho parlato* e ("and") *le ho chiesto; mi dà fastidio che > a me mi dà fastidio che; devo dire ci sono > devo dire che* ("that") *ci sono*. The far commonest addition is that of the conjunction *e* ("and"), that occurs in nearly half of the cases (45%), followed by articles 15.4% (6), adverbs 15.4% (6), pronouns 10.3% (4) (*cf.* Table 6).

	Frequency	*Percent*
noun	2	5.1
adverb	6	15.4
adjective	1	2.6
conjunction	18	46.2
pronoun	4	10.3
article	6	15.4
preposition	2	5.1
Total	*39*	*100.0*

Table 6 - Grammatical category of additions

Additions generally affect function words (in 72.5% of the cases, with conjunctions – *e* – and articles – *la* – inserted in the textual material), while lexical units are added in 25% of the errors of this kind. While from the semantic point of view additions rarely change utterance meaning. Meaning is preserved in 90% of the cases, and partially preserved in 7.5%.

Among deletion, elision of one or more elements from transcriptions, is common misdetection of repetitions (21.6% of deletion cases), especially of function words not playing any role other than fillers (*fa la parte di quello che mi prende in giro*, instead of *che che*). Examples of deletions are: i*n economia non* ("not") *sono tranquillo > in economia sono tranquillo; per restituire all'Italia > per l'Italia; è arrivato anche dal ministero > è arrivato dal ministero; e che sono meno inquinanti > e meno inquinanti.* Deletions occur in 43.7% of errors (199 cases).

Deletions occur mainly at the lexical/syntactic 29.1% (58), in repetition 21.6% (43), at lexical level 19.1% (38), and in function words 16.6% (33) (*cf.* Table 7). At lexical/syntactic level, more than one word is involved in deletion (entire phrases often), and consequences for the overall understanding result menaced. On the contrary in the case of repetition (of words, constituents, fillers) no effect on meaning is caused by cancellation. Deletions often regard entire constituents (41.7% of cases), and are generally more dangerous for meaning preservation: 50.3% of cases are not affected semantically by the error, while 16.6% are partially affected and 33.2% lead to misunderstanding. Misunderstanding mainly occurs when more than one word is omitted (at lexical/syntactic level).

	Frequency	*Percent*
lexical	38	19.1
function word	33	16.6
syntactic	22	11.1
phatic expression	3	1.5
lexical/syntactic	58	29.1
replanning deleted	2	1.0
repetition deleted	43	21.6
Total	*199*	*100.0*

Table 7 - Error types in deletion

Grammatical categories of deleted words (when single words are involved, 80 cases, 40.2%) are adverbs 22.5% (18), verbs 20% (16), conjunctions 16.3% (13), pronouns 15.0% (12), prepositions 8.8% (7), nouns 7.5% (6), adjectives 7.5% (6), articles 2.5% (2).

Movement, where one or more elements are misplaced in the utterance order sequence, is the least frequent phenomenon with only 2.4% of total error occurrences, in just 11 cases. Movement rarely changes the overall meaning of the utterance (18.2% of movement cases), and always involves entire sentence fragments and not single words (*sull'appennino centrale e sul medio versante* instead of *sul medio versante e sull'appennino centrale*). Examples of movement errors are: *otto casi finora > finora otto casi; è esattamente quello di > esattamente è quello di; anche se parlano troppo > se anche parlano troppo; delle mamme delle nonne > delle nonne delle mamme.*

4. Errors in spoken corpora transcripts

The recently released CLIPS corpus (*Corpora e Lessici di Italiano Parlato e Scritto*) of spoken Italian has been used to test different versions of the transcripts of the audio material with different revisions and the final version in order to observe error typologies regarding the mere transcription of words (thus excluding phenomena such as pauses, other vocal behaviour, noises, etc.). The corpus subsection that has been analyzed is media subcorpus including radio and television broadcasts regarding four typologies: entertainment, news, culture, advertisement. The media subcorpus consists of 50% radio broadcasts and 50% television broadcasts, from national and local networks for a total amount of sixteen hours of recordings. Of these sixteen hours only a portion of the corpus has been transcribed: about sixty minutes of national broadcasts (thirty minutes of radio and thirty of television) and about eighteen minutes for each of the fifteen cities where the recordings took place (Bari, Bergamo, Bologna, Cagliari, Catanzaro, Florence, Genoa, Lecce, Milan, Naples, Palermo, Parma, Perugia, Rome, Venice), for a total of 330 minutes (5.5 hours).

The corpus has been processed at different levels. Orthographic transcripts have been produced by different transcribers (a total of 29 transcribers for the whole 100 hours corpus) and subsequently revised by different researchers. A smaller section of the corpus has also been phonetically annotated, thus leading to a further revision of the full transcripts. No explicit trace of the number of revisions are given in the public documentation.

Recordings have been analyzed and compared to the basic orthographic transcription in order to understand patterns in errors in detecting the speech flow. The transcripts have been analyzed in the public version available on the CLIPS website (www.clips.unina.it, last accessed 8 July 2007) . The orthographic transcripts follow a guidelines document (Savy, 2007) giving detailed information about conventions to be used in the transcripts, such as main objectives, file formats and names, header information, orthographic conventions, *etc*. Savy (2007: 2) proposes a distinction between the transcription as the basic coding or *representation* of speech and the *interpretation* process that is involved in annotating the transcript with additional information. She further claims that the basic transcription as representation, compared to the mark-up and additional annotation, is "at the lowest level of complexity" (2007: 4). On the contrary, as it will be showed in the following paragraphs the representation of the mere sequence of lexical items spoken requires a large amount of interpretation that often emerges in the deletion, addition, substitution or movement of the items in the transcribed sequence.

Minutes of recordings	330
Errors reported	135
Avrg. nr. errors per minute	0.41

Table 8 - Overview of errors in the CLIPS media

Compared to the experimental data the number of errors in corpus data are definitely smaller. While in experimental data the average number of errors per minute of listening (without counting repetitions) is 11.38, in corpus data one error appears roughly every two minutes (0.41, *cf.* Table 8). Considering the fact that transcription

in corpus data is self-administered, that the transcriber can listen to the sequence the number of times that he/she considers appropriate and that the transcript is followed by revision, still the number of errors appears high.

If we look at general change the most common error typology is substitution (60 cases, 44.4%), and deletions (54 cases, 40%), while cases of addition (18 cases, 13.3%) and movement (3 cases, 2.2%) are uncommon (*cf.* Figure 2). In 83.7% of the cases the error affects only single words, while in 15.6% it affects entire sentences or phrases.

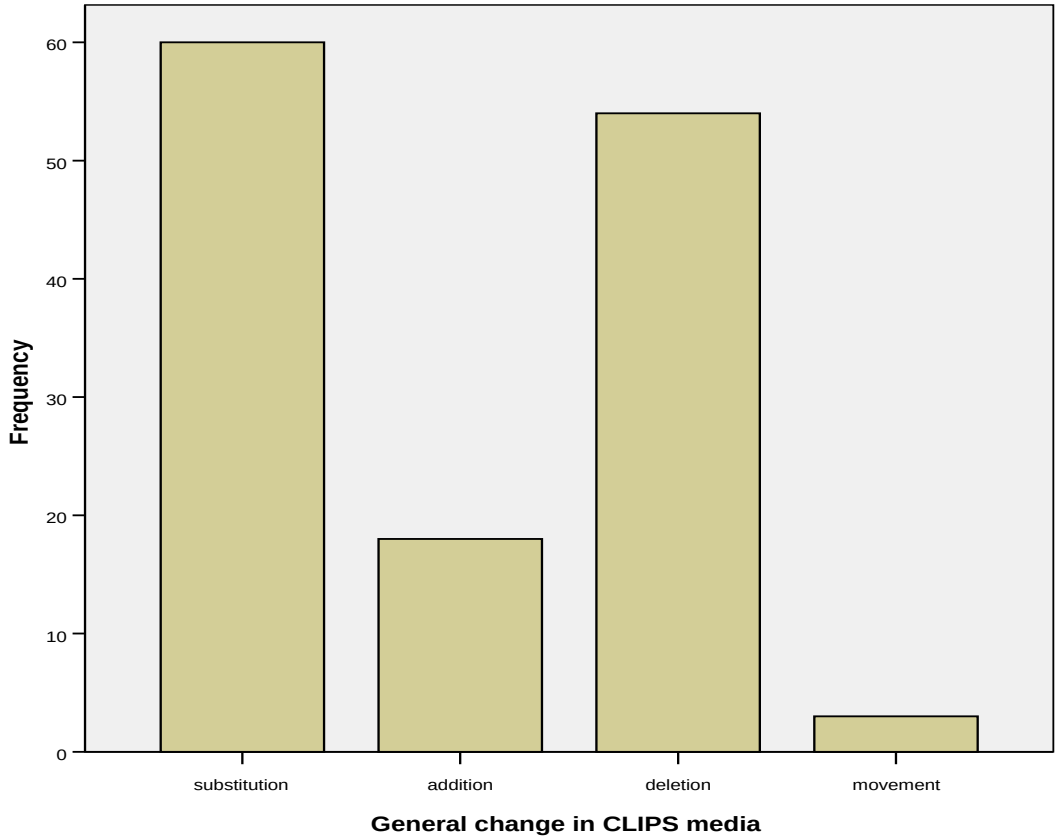

Figure 2 - Type of change in CLIPS media (corpus data)

Relative frequencies perfectly match experimental data where proportions of general error typology are exactly the same. If we look at meaning preservation in errors in corpus data (*cf.* Table 9), relative frequencies tend to distribute in the same way as in experimental data (*cf.* Table 3). Cases of meaning preservation are such as *l'iscrizione sul* ("on") *registro degli indagati* is transcribed as > *l'iscrizione nel* ("in") *registro degli indagati* or *nei suoi riguardi* is transcribed as > *nei suoi confronti*.

	Frequency	*%*
yes	67	49,6
partial	21	15,6
no	47	34,8
Total	*135*	*100,0*

Table 9 - Meaning preservation in CLIPS media

Errors distribute evenly among the diamesic variation between radio and television broadcasts, with a slight prevalence for the television section (*cf.* Table 10).

	Frequency	*Percent*
Radio	66	48,9
Television	69	51,1
Total	*135*	*100,0*

Table 10 - Diamesic variation in CLIPS media errors

If we look more closely to finer error typologies we observe a definite prevalence for the suppression of units of speech, with 55 cases (40.7%), followed by lexical substitutions (42 cases, 31.1%, *cf.* Table 11).

	Frequency	*Percent*
lexical switch	42	31,1
sing/plur switch	5	3,7
function word substitution	6	4,4
insertion of words	17	12,6
missing words	55	40,7
syntactic misplacement	3	2,2
verb conjugation error	1	,7
phonetic variant	6	4,4
Total	*135*	*100,0*

Table 11 - Error typologies in CLIPS media

5. Conclusions

The present research was both meant to provide hints on human understanding and creative repair in a linguistic re-production task and suggest specific error typologies that can and do occur in linguistic corpora transcription and that are not easily detectable in automatic post-editing procedures without direct access to the spoken audio material. The most striking finding regards the amount of repair that does not rely of linguistic form but on creative unconscious reconstruction made by the transcriber, that generally tends to preserve utterance meaning. The transcriber attributes intentions and beliefs to the voice heard, and tends to filter inevitably the spoken sounds re-interpreting them in a way that is always both grammatical and meaningful. This re-interpretation often does not lead to a general misunderstanding of what was heard but to a reformulation of a single portion of the speech flow.

In a sense, the listener is never a reliable listener, unless its main task is meaning-centred. Even when explicitly asked (and trained) to concentrate on form (and on the sequence of exact words to reproduce), his attitude turns toward meaning-centred practices. A possible interpretation of this findings might be that ordinary understanding behaviour is strictly focused on meaning rather than form, so that, even with the best possible audio quality, when trying to concentrate attention on the reconstruction of linguistic form, we tend to shift and rely on our understanding strategies, that lead us to re-create text in a plausible way.

There are some apparent similarities among transcription errors and the so-called *slips of the ear, Verhören* or *lapsus auris* (Bond, 1999; Chiari, 2005; Voss, 1984), intended as misunderstanding of the perceived sound chain that leads to reinterpretation. There are common detecting problems when dealing with slips of the ear since there is no direct access to the listeners' understanding process, unless the listener himself signals, by repetition for example, his suspected misunderstanding (whether conscious or unconscious). In addition to some relevant methodological problems common to slips of the tongue (Ferber, 1991) generally slips of the ear are characterized by an accepted divergence in meaning reconstruction from the original utterance produced by the speaker. In transcription errors, however, the mistake cannot be attributed to actual misunderstanding but to subsequent interventions relying on what the hearer has actually understood (and well understood most of the times).

Even though structural changes (additions, deletions, inversions and substitutions) might seem similar to those occurring in slips of the tongue and slips of the ear it is not possible at the moment to investigate the subject in a through way, since corpora of detected slips of the ear are not currently available for Italian. It is interesting to note that corpus and experimental data tend to agree in the relative frequencies distribution of structural changes. An interesting further object of exploration should be the evaluation of frequencies of those changes in slips of the ear. Movement for example, that corresponds to the slip typology of inversion is extremely rare in transcription errors, while in slips is a quite common surface category.

The central theoretical point regarding transcription errors lies on listeners' repair strategies. Repair, which is obviously not perceived as such by the transcriber, can be due to different grounds. The conversion from speech to writing certainly plays a role in the adaptation of the spoken chain to a more consistent form which exhibits more explicit cohesive markers (deletion of repetition, especially those representing hesitation or insertions of the *e* "and" coordination), as well as error correction (agreement reconstruction, or the redundant expression *a me mi dispiace* becoming for the transcriber *a me dispiace*). Attention factors probably play a role as well as memory spans, especially in experimental settings where the participant cannot control the administration of the audio repetitions.

A further point regards error patters and repairs that suggest that there might be weak elements in a spoken discourse which are more often subject to deletion or repair during transcription. In a number of utterances included in the experiment different transcribers were submitted the same utterance and made the same errors, such as *E un quasi decalogo di consigli pratici è arrivato* anche *dal ministero delle attività produttive* ("and quasi-decalogue of practical advices has been provided also by the production activities ministry"), where *anche* ("also") is systematically deleted.

The presence of an error (especially those that imply substitution of verb tense or person, and singular/plural switching) often produces the occurrence of other errors in the following words, since the transcriber tends to repair textual cohesion signals. For example, since the transcriber has erroneously perceived a singular subject (*il corridore*) in the utterance (*I soccorritori avrebbero avuto problemi*), the rest is conjugated with a verb agreement in the singular form (*avrebbe avuto problemi*).

Finally, there are some cases in which the transcriber is faced with tasks which are more complex than ordinary. The presence of noise and superimpositions, homophonous sequences, problems of word boundary identification, transcriptions of linguistics errors (at different levels, grammatical, phonetic) or of speech characterized by greater unpredictability in form or content (language disorders, psychopathologies, extremely old or young speakers, etc.). These cases pose further questions. Corpus data used in the present study did not show extremes which could actually be included in these cases, but research on understanding for transcription practices could benefit greatly from an investigation covering more erratic speech input.

Better knowledge of transcription errors allows improved planning of instruction guidelines supplied to transcribers (training the ears and training the mind towards formal and superficial linguistic elements) and improvement in the correction and revision phases during corpus processing and annotation. Nevertheless, even trained transcribers tend to make mistakes of which they remain unaware.

References

Ashmore, M. & Reed, D. (2000) 'Innocence and Nostalgia in Conversation Analysis: The Dynamic Relations of Tape and Transcript'. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research* 1(3), Available on-line from: http://www.qualitative-research.net/fqs-texte/3-00/03-00ashmorereed-e.htm (accessed: 28 June 2007).

Bond, Z. S. (1999) Slips of the ear. San Diego, Calif: Academic Press

Chafe, W. (1995) Adequacy, user-friendliness, and practicality in transcribing, in G. Leech, G. Myers & J. Thomas (eds.), Spoken English on computer:Transcription, mark-up, and application. Harlow, England: Longman.

Chiari, I. (2005) Condizioni, limiti e analogie del lapsus linguae e del lapsus auris, in A. Frigerio & S. Raynaud (eds.), Significare e comprendere: la semantica del linguaggio verbale. Roma: Aracne.

Chiari, I. (2006a) 'Slips and errors in spoken data transcription'. Paper presented at the Proceedings of 5th International Conference on Language Resources and Evaluation LREC2006, Genova. Genova: CD-ROM, ELDA 1596-1599.

Chiari, I. (2006b) 'Spoken corpora and transcription errors/ Звучащие корпуса и ошибки транскрипции'. Paper presented at the Труды Международной научной конференции «Корпусная лингвистика 2006» / Proceedings of the International Conference «Corpus Linguistics 2006», СПб. СПб: Изд-во С.-Петерб. ун-та Изд-во С.-Петерб. ун-та.

Connell, D. C. & Kowal, S. (1999) 'Transcription and the Issue of Standardization'. *Journal of Psycholinguistic Research* 28, 103-120.

Cook, G. (1995) Theoretical issues: transcribing the untranscribable, in G. Leech, G. Myers & J. Thomas (eds.), Spoken English on Computer: Transcription, Markup and Applications. Harlow: Longman.

Corder, S. P. (1973) Introducing applied linguistics. Harmondsworth: Penguin Education.

Derville, B. (1997) 'Expression orale: du bon usage des transcriptions'. *Le Francais dans le Monde* 287(Feb-Mar), 29-32.

Edwards, J. A. & Lampert, M. D. (1993) Talking data : transcription and coding in discourse research. Hillsdale, N.J.: Lawrence Erlbaum Associates.

Ferber, R. (1991) 'Slip of the tongue or slip of the ear? On the perception and transcription of naturalistic slips of the tongue'. *Journal of Psycholinguistic Research* 20(2), 105-122.

Lapadat, J. C. (2000) 'Problematizing transcription: purpose, paradigm and quality'. *International Journal of Social Research Methodology* 3, 203-219.

Leech, G. N., Myers, G. & Thomas, J. (1995) Spoken English on computer : transcription, mark-up, and application. Harlow, essex, England ; New York: Longman.

Lindsay, J. & O'Connell, D. (1995) 'How do transcribers deal with audio recordings of spoken discourse?' *Journal of Psycholinguistic Research* 24(2), 101-115.

Ochs, E. (1979) Transcription as theory, in E. Ochs & B. Schieffelin (eds.), Developmental pragmatics. New York: Academic Press.

Pallaud, B. (2002) 'Erreurs d'ecoute dans la transcription de donnees orales'. *Revue PArole* 22-23-24, 267-229.

Pallaud, B. (2003) 'Splendeurs et misères de la transcription'. *Cliniques méditerranéennes* 68(2), 59-73.

Powers, W. R. (2005) Transcription techniques for the spoken word. Lanham, MD: AltaMira Press.

Romero, C., O'Connell, D. C. & Kowal, S. (2002) 'Notation Systems for Transcription: An Empirical Investigation'. *Journal of Psycholinguistic Research* 31(6), 619-631.

Savy, R. (2007) Specifiche per la trascrizione ortografica annotata dei testi raccolti (CLIPS/W1-a4/PTA/001/ver3).

Vigouroux, C. B. (2007) 'Trans-scription as a social activity: An ethnographic approach'. *Ethnography* 8(1), 61-97.

Voss, B. (1984) Slips of the Ear: Investigations Into the Speech Perception Behaviour of German Speakers of English. Gunter Narr Verlag.