

## Investigation about speech recognition (January - May 2006) report of Gian Paolo Trivulzio

Our President, Mister Jaroslav Zaviacic, suggested on 20<sup>th</sup> December 2005, to investigate about the accuracy of speech recognition, giving some data raising from his experience in Czech Republic.

I include here below his message

----- Original Message -----

**From:** [Jaroslav Zaviacic](#)

**To:** ['Trivulzio Tin'](#) ; ['Cees van Beurden'](#)

**Cc:** ['Zaviacic Jaroslav Pres. Intersteno'](#) ; ['Panzer Mauro Jury'](#) ; ['Danny Devriendt'](#) ; ['Ramondelli Fausto Senato'](#) ; ['Jorge Alberto Bravo'](#) ; [neubauer@FH-Aachen.de](mailto:neubauer@FH-Aachen.de) ; ['David J. Rogala'](#)

**Sent:** Tuesday, December 20, 2005 12:39 AM

**Subject:** RE: Experiences of Intersteno in speech recognition

Lieber Gian Paolo,

danke für Überblick zur Spracherkennung: es zeigt, dass bei dem Wettbewerb "Schnelle Übertragung" - wo die Geschwindigkeit, Perfektion der Übertragung, aber auch die Übertragungszeit ausgewogene Kriterien für Ergebnis vorstellen, lässt sich jede Technik zulassen, sogar kombiniert mit der Tonaufnahme. Damit zeigt es sich, wie nicht nur der Wetschreiber, sondern auch von ihm gewählte Technik effektiv ist.

Trotzdem wäre es höchst objektiv auch zu testen, wieviel Prozent des fortlaufenden Textes bei der Spracherkennung direkt (ohne nachträglichen Korrekturen) richtig übertragen wird. Auf Tschechisch nur etwa 70 %, wenn Herr Verusso nach 8 Minuten "Umdiktierens" anschließend 19 Minuten lang es verbessert (dazu noch mit der Tonaufnahme), kann es auch mit 70 - 80 % auf Italienisch stimmen.

Freundlich aus Prag

Jaroslav

This message followed a previous one sent on 14<sup>th</sup> December as here below:

----- Original Message -----

**From:** [ZAV](#)

**To:** ['Trivulzio Tin'](#) ; ['Zaviacic Jaroslav Pres. Intersteno'](#) ; ['Ramondelli Fausto Senato'](#) ; ['Cees van Beurden'](#) ; ['Panzer Mauro Jury'](#)

**Cc:** [neubauer@FH-Aachen.de](mailto:neubauer@FH-Aachen.de) ; ['Jorge Alberto Bravo'](#) ; ['Rogala David \(el. adresa\)'](#) ; ['Devriendt Danny \(el. adresa\)'](#) ; ['Sager Franz \(el. adresa\)'](#) ; ['Schindelasch Hannelore \(el. adresa\)'](#) ; ['Dierks Waltraut \(el. adresa\)'](#)

**Sent:** Wednesday, December 14, 2005 12:32 AM

**Subject:** RE: Reporting in the German courts with speech recognition - second batch of 1300 installations.

Lieber Gian Paolo, liebe Freunde,

die Information "Reporting in the German courts ..." zeigt die Wichtigkeit meines Vorschlages die Spracherkennung als weitere Technik bei unserem Wettbewerb "Schnelle Übertragung" zulassen, die Technik der Stenomask mit Tonaufnahme bei Stenowettbewerben verbieten, aber paralell dazu

eine neue Disziplin "Reporting" einführen, bei der die Übertragungszeit zB in der Länge der Ansagezeit von einem Stenodiktat erlaubt wird, damit objektiv verschiedene Techniken der Textfassung bei gesprochenem Wort (Diktat) verglichen werden.

Ohne Beschränkung der Übertragungszeit haben alle Ergebniszahlen nur die Aussagekraft, wie schnell man diktieren kann, nicht wie schnell die Spracherkennung arbeitet.

Ich schlage vor in Deutschland einen Test-Wettbewerb zu organisieren, damit auch Angaben aus der Mitteilung "Reporting in the German courts ..." objektiv überprüft werden: "... Dragon NaturallySpeaking Legal converts speech into text at up to 160 words per minute, and is tailored to the special workflow and vocabulary used in the legal sector. Dragon NaturallySpeaking Legal assures unsurpassed speech recognition accuracy - up to 99 percent - thanks to its standard vocabulary of more than 400,000 terms."

Nach 11 Jahren eine Arbeitsgruppe aus der Technischen Hochschule Liberec mit 350 000 eingetragenen Wörtern kommt zu ähnlichen Ergebnissen, die auch ich notiert habe: Geschwindigkeit kaum 40 Wörter pro Minute, Perfektion um 70 % bei fortlaufendem Text, etwa 90 % bei einigen Textmodulen.

Mag sein, die niedrigen Zahlen mit der komplizierten slawischen Sprache hängen zusammen. Deswegen sollten wir solchen Test- bzw. Versuchswettbewerb in deutscher Sprache initiieren, bei dem ein Team Gerichtsspezialisten und ein Team von Stenografenverband / Intersteno Deutschland sich trifft und durch ein Intersteno-Test-Wettbewerb die objektive Zahlen über Geschwindigkeit und Perfektion von mehreren Techniken erfährt.

Freundliche Grüße aus Prag

Jaroslav Zaviacic

-----Original Message-----

**From:** Trivulzio Tin [mailto:gianpaolo.trivulzio@tin.it]

**Sent:** Tuesday, December 13, 2005 6:21 PM

**To:** Zaviacic Jaroslav Pres. Intersteno; Ramondelli Fausto Senato; Cees van Beurden; Panzera Mauro Jury

**Cc:** neubauer@FH-Aachen.de; Jorge Alberto Bravo; Crippa Corti Maria Luisa Delegata Italia

**Subject:** Reporting in the German courts with speech recognition - second batch of 1300 installations.

Following this invitation we have set-up a procedure for collecting information and making tests when opportunity arose, this procedure foresees two steps (up to now only the first step has been used):

1 - reading of a written text for 1 minutes and calculating number of words, characters and faults. This activity is comparable to that of text production in Intersteno competitions.

2 - 3 minutes dictation at 100, 110, 120 Italian words, equivalent to about 250, 275, 300 syllables per minute. This is a typical 'steno competition'. Up to now only two persons has done this test, so the pattern is too low; results will be sent as soon as we will be able to have at least 10 tests.

Counting of errors is done without considering the reason which originated the fault, i.e. if the persons says a word different from what is written, it is in any case considered a fault. The **purpose is not to test the speech engine, but to test the**

**overall process with a speaker having at least a minimal training on speech recognition.**

Collected results (up to now), are shown in the table at the next pages.

Some general information and remarks

1 - all attendants used open microphones, many of them using external Andrea USB pod

2 - the age of participants (here not included for privacy reason) varies from 27 till 49.

3 - All participants are daily using computer for text production - Internet etc.

4 - Two texts were used (which are included at the end of this document) having more or less the same linguistic level of difficulty in reading.

1. Text 1 - needs only the basic training on how to speak to a computer - the basic adaptation of software to the voice of the speaker and the voice command for the punctuation signs.
2. Text 2 - in addition to the previous, requires a good mastery of voice commands (new paragraph - letters with capital like 'parco', which is usually written with lower case, but in this text is referred to a specific 'parco' so it is written with capital letter 'Parco').

5 - the calculation of faults is made according to the Intersteno standard for keyboarding i.e. 1 fault per word independently of the quantity of wrong letters. With speech recognition all words in the dictionary are correctly written from the orthographic point of view.

6 - calculation of percentage is related as a ratio between the number of faults and the total number of words (Italian words of course). For a better evaluation the quantity of words is also indicated in syllables (true figures not a standard conversion), as well as in number of characters (also here they are true figures).

7 - The ratio of syllables per word and characters per word is also included for additional statistical comparison.

I hope that our German friend will release similar info of their research, so we can have a wider picture.

*Gian Paolo Trivulzio*

Milano, 9<sup>th</sup> June 2006

Personal information					Test results and statistical info								Hardware and Software			
N.	activity	sex	use of speech software	speech recognition learning	test	speed (words)	faults	accuracy %ge	speed (syllables)	speed (char.s)	syllable/words	characters/words	Computer	Memory (Mega)	Speech Software	Editor
1	reporter	m	random/3years	self-studium	1	136	5	96,32	319	894	2,35	6,57	laptop	256	Via Voice	Word
2	reporter	f	random/1year	self-studium	1	144	4	97,22	338	947	2,35	6,58	lap top	256	Via Voice	Word
3	secretary	f	beginner	training course – 3 hours	1	92	3	96,74	220	608	2,39	6,61	laptop/Andrea Usb Pod	516	Dragon 7 professional	Word
4	reporter	f	beginner	training course – 6 hours	1	114	3	97,37	274	756	2,40	6,63	Desktop/Andrea Usb Pod	516	Dragon 7 professional	Word
5	reporter	f	beginner	training course 6 hours	1	104	4	96,15	246	678	2,37	6,52	Desktop/Andrea Usb Pod	516	Dragon 7 professional	Word
6	reporter	f	beginner	training course 3 hours	1	132	5	96,21	319	880	2,42	6,67	desktop/Andrea Usb Pod	516	Dragon 7 professional	Word
7	reporter	f	beginner	training course 3 hours	1	123	2	98,37	293	813	2,38	6,61	desktop/Andrea Usb Pod	516	Dragon 7 professional	Dragon Pad
8	text producer	f	beginner	training course 12 hours	2	105	3	97,14	253	695	2,41	6,62	desktop/Andrea Usb pod	516	Dragon 7 professional	Word
9	text producer	f	beginner	training course 12 hours	2	118	4	96,61	274	758	2,32	6,42	Desktop/Andrea Usb pod	516	Dragon 7 professional	Word

10	text producer	f	beginner	training course 12 hours	2	106	2	98,11	257	706	2,42	6,66	desktop/Andrea Usb pod	516	Dragon 7 professional	Word
11	text producer	f	beginner	training course 12 hours	2	133	5	96,24	307	854	2,31	6,42	Desktop/Andrea Usb pod	516	Dragon 7 professional	Word

## Test 1

### Test 1

In quasi tutte le grandi città degli stati industrializzati occidentali si delineano sviluppi che nei prossimi anni potrebbero modificare in modo significativo le loro città e anche le condizioni di vita degli abitanti. In parte ciò porta al fatto che la politica municipale di queste città verrà posta di fronte a problemi completamente nuovi. Quando in parecchi paesi si è dato inizio ai lavori sullo sviluppo della pianificazione urbanistica, non si sono registrati ovunque consensi entusiastici. Ci sono state alcune opinioni contrarie. Più volte si è riscontrato un chiaro disinteresse. Indagando sulle cause di questo mancato interesse e più precisamente sul rifiuto, si è riscontrata l'opinione diffusa che il momento attuale sembra davvero inadatto per un piano di sviluppo cittadino. Soprattutto in tempi di serie difficoltà e problemi, soprattutto in ambito economico, ci si dovrebbe concentrare per risolvere i problemi più urgenti della politica quotidiana e non effettuare ricerche approfondite sullo sviluppo a lungo termine delle basi della vita della città.

### Counting of syllables

In/ qua/si /tut/te/ le /gran/di /cit/tà /de/gli /sta/ti/ in/du/stria/liz/za/ti/  
oc/ci/den/ta/li /si /de/li/nea/no /svi/lup/pi /che /nei /pros/si/mi/ an/ni  
/po/treb/be/ro /mo/di/fi/ca/re/ in/ mo/do /si/gni/fi/ca/ti/vo/ le/ lo/ro /cit/tà/ e/  
an/che/ le /"con/di/zio/ni"/ di /vi/ta /de/gli/ a/bi/tan/ti./ In/ par/te /ciò /por/ta/  
al/ fat/to /che/ la /po/li/ti/ca /mu/ni/ci/pa/le /di /que/ste /cit/tà /ver/rà /po/sta /di  
/fron/te/ a /pro/ble/mi /com/ple/ta/men/te /nuo/vi./ Quan/do/ in/ pa/rec/chi  
/pae/si /si/ /è /da/to/ i/ni/zio/ /ai/ la/vo/ri /sul/lo /svi/lup/po /del/la

/pia/ni/fi/ca/zio/ne/ ur/ba/ni/sti/ca,/ non/ si /so/no/ re/gi/stra/ti/ o/vun/que  
/con/sen/si/ en/tu/sia/sti/ci./ Ci /so/no /sta/te/ al/cu/ne/ o/pi/nio/ni  
/con/tra/rie./ Più /vol/te /si/ /è/ ri/scon/tra/to/ un/ chia/ro /di/sin/te/res/se./  
In/da/gan/do /sul/le /cau/se /di /que/sto /man/ca/to/ in/te/res/se/ e /più  
/pre/ci/sa/men/te /sul/ ri/fiu/to,/ si/ /è/ ri/scon/tra/ta/ l'o/pi/nio/ne /dif/fu/sa  
/che/ il/ mo/men/to/ at/tua/le /sem/bra /dav/ve/ro /"i/na/dat/to"/ per/ un/  
pia/no /di /svi/lup/po /cit/ta/di/no./ So/prat/tut/to/ in/ tem/pi /di /se/rie  
/dif/fi/col/tà/ e /pro/ble/mi,/ so/prat/tut/to/ in/ am/bi/to/ e/co/no/mi/co,/ ci /si  
/do/vreb/be /con/cen/tra/re /per/ ri/sol/ve/re/ i /pro/ble/mi /più/ ur/gen/ti /del/la  
/po/li/ti/ca /quo/ti/dia/na/ e /non/ ef/fet/tua/re/ ri/cer/che/ ap/pro/fon/di/te  
/sul/lo /svi/lup/po/ a/ lun/go /ter/mi/ne /del/le /ba/si /del/la /vi/ta /del/la /cit/tà./

**Le sillabe sono = 386. Le parole sono: 179 media 2.156 sillabe per  
parola**

## Test 2 –

**Leggere tranquillamente per un minuto, dopo il via.**

### **Imposta carattere Arial 20**

Il parco naturale ha investito molte energie durante l'estate ed ha da poco concluso numerose attività rivolte ad un ampio pubblico, registrando la partecipazione diretta di turisti e residenti. Inoltre il Parco ha tessuto una fitta rete di collaborazioni con diversi soggetti che operano sul territorio, in ambiti di promozione turistica e divulgazione, in ambito culturale, artistico e scientifico, nel settore della produzione di tipicità alimentari e di servizi logistici. L'obiettivo che il Parco Naturale si propone è quello di sollecitare la conoscenza, il rispetto e quindi la tutela dell'ambiente.

È fondamentale che si impari a curiosare con discrezione tra le cose della natura: per conoscere con la mente e con il cuore, per vedere il mondo da punti di vista diversi, con gli occhi degli altri, cercando di superare i pregiudizi culturali che ogni persona porta con sé. Al fine di perseguire tale alto obiettivo, sono state studiate svariate proposte, dando così la possibilità a residenti e turisti di vivere appieno un periodo di vacanza nel territorio del Parco.

Per un'estate da Parco è l'ombrellone che ha tenuto sotto di sé le svariate e coinvolgenti attività proposte in passato. Dopo una giornata di sole trascorsa nelle suggestive valli del Parco, è stato ammirato, tra i monti, il tramonto che emoziona quando, all'arrivo della notte, si sentono i versi degli animali notturni, mentre noi camminiamo lungo i sentieri del bosco! Non potevamo certo perderci anche la piacevole lettura del cielo stellato prima di addormentarci fra le braccia di Morfeo! Non stiamo confondendo poesia e realtà. L'estate ha permesso di immergersi e vivere il Parco in ogni momento della giornata, dall'alba al tramonto, sotto il cielo stellato e nel bosco di notte. Con la collaborazione delle Aziende per il Turismo abbiamo aspettato il sorgere dell'alba assaporando pane nero e marmellata di mirtilli ed ammirato i colori crepuscolari tra le cime rosate delle Dolomiti.

Il/ par/co /na/tu/ra/le/ ha/ in/ve/sti/to /mol/te/ e/ner/gie /du/ran/te /l'e/sta/te/ ed/ ha/ da /po/co /con/clu/se/ nu/me/ro/se/ at/ti/vi/tà /ri/vol/te/ ad/ un/ am/pio /pub/bli/co,/ re/gi/stran/do /la /par/te/ci/pa/zio/ne /di/ret/ta /di /tu/ri/sti/ e /re/si/den/ti./ Inol/tre/ il/ Par/co/ ha /tes/su/to/ u/na /fit/ta /re/te /di /col/la/bo/ra/zio/ni /con/ di/ver/si /sog/get/ti /che/ o/pe/ra/no /sul/ ter/ri/to/rio,/ in/ am/bi/ti /di /pro/mo/zio/ne /tu/ri/sti/ca/ e



/di/vul/ga/zio/ne,/ in/ am/bi/to /cul/tu/ra/le,/ ar/ti/sti/co/ e /scien/ti/fi/co,/ nel/ set/to/re  
 /del/la /pro/du/zio/ne /di /ti/pi/ci/tà/ a/li/men/ta/ri/ e /di /ser/vi/zi /lo/gi/sti/ci./  
 L'o/biet/ti/vo /che/ il/ Par/co /Na/tu/ra/le /si /pro/po/ne/ /è /quel/lo /di /sol/le/ci/ta/re /la  
 /co/no/scen/za,/ il/ ri/spet/to/ e /quin/di /la /tu/te/la /del/l'am/bien/te./  
 È fon/da/ment/a/le /che /si/ im/pa/ri/ a /cu/rio/sa/re /con/ di/scre/zio/ne /tra /le /co/se  
 /del/la /na/tu/ra:/ per/ co/no/sce/re /con/ la /men/te/ e /con/ il/ cuo/re,/ per/ ve/de/re/ il/  
 mon/do /da /pun/ti /di /vi/sta /di/ver/si,/ con/ gli/ oc/chi /de/gli/ al/tri,/ cer/can/do /di  
 /su/pe/ra/re/ i /pre/giu/di/zi /cul/tu/ra/li /che/ o/gni /per/so/na /por/ta /con/ sé./ Al/ fi/ne  
 /di /per/se/gui/re /ta/le/ al/to/ o/biet/ti/vo,/ so/no /sta/te /stu/dia/te /sva/ria/te /pro/po/ste,  
 dan/do /co/sì /la /pos/si/bi/li/tà/ a /re/si/den/ti/ e /tu/ri/sti /di /vi/ve/re/ ap/pie/no/ un/  
 pe/rrio/do /di /va/can/za /nel/ ter/ri/to/rio /del/ Par/co./  
 Per/ un/'e/sta/te /da /Par/co/ /è /l'om/brel/lo/ne /che/ ha /te/nu/to /sot/to /di /sé /le  
 /sva/ria/te/ e /coin/vol/gen/ti/ at/ti/vi/tà /pro/po/ste/ in/ pas/sa/to./ Do/po/ u/na  
 /gior/na/ta /di /so/le /tra/scor/sa /nel/le /sug/ge/sti/ve /val/li /del/ Par/co,/ è /sta/to/  
 am/mi/ra/to,/ tra/ i /mon/ti,/ il/ tra/mon/to /che/ e/mo/zio/na /quan/do,/ al/l'ar/ri/vo  
 /del/la /not/te,/ si /sen/to/no/ i /ver/si /de/gli/ a/ni/ma/li /not/tur/ni,/ men/tre /noi  
 /cam/mi/nia/mo /lun/go/ i /sen/tie/ri /del/ bo/sco!/ Non/ po/te/va/mo /cer/to /per/der/ci/  
 an/che /la /pia/ce/vo/le /let/tu/ra /del/ cie/lo /stel/la/to /pri/ma /di/ ad/dor/men/tar/ci /fra /le  
 /brac/cia /di /Mor/feo!/ Non/ stia/mo /con/fon/den/do /poe/sia/ e /real/tà./ L'e/sta/te/ ha  
 /per/mes/so /di/ im/mer/ger/si/ e /vi/ve/re/ il/ Par/co/ in/ o/gni /mo/men/to /del/la  
 /gior/na/ta,/ dal/l'al/ba/ al/ tra/mon/to,/ sot/to/ il/ cie/lo /stel/la/to/ e /nel/ bo/sco /di  
 /not/te./ Con/ la /col/la/bo/ra/zio/ne /del/le/ /A/zien/de /per/ il/ Tu/ri/smo/ ab/bia/mo/  
 a/spet/ta/to/ il/ sor/ge/re /del/l'al/ba/ as/sa/po/ran/do /pa/ne /ne/ro/ e /mar/mel/la/ta /di  
 /mir/til/li/ ed/ am/mi/ra/to/ i /co/lo/ri /cre/pu/sco/la/ri /tra /le /ci/me /ro/sa/te /del/le  
 /Do/lo/mi/ti./

**Le sillabe sono = 700. Le parole sono: 312 media 02,2436 sillabe per parola**