# Relationship between Workload and Duration of ATC Voice Communications


Natalia Rodríguez Uclés
PhD Candidate – Researcher
CRIDA A.I.E, ATM R&D Reference Center
Madrid, Spain
nrucles@e-crida.aena.es

José Manuel Cordero García
Principal Researcher
CRIDA A.I.E, ATM R&D Reference Center
Madrid, Spain
jmcordero@e-crida.aena.es



*Abstract*—**This paper establishes the relationship between Air Traffic Controller Officer voice communications and workload, across various operative sectors in different times of the day, composing a wide and representative sample for the purposes of the study. It is generally assumed that Air Traffic Control system complexity is coupled to operational control difficulty and thus to Workload. In order to estimate controller workload, several complex and subjective factors are taken into account: experience, capabilities, age, motivation, mood and operational problem resolution strategies among the most relevant ones. Additionally, workload estimation is also influenced by other Air Traffic Control system parameters such as traffic volume, operational restrictions, weather, unexpected events or voice communications. Past studies in the area of voice communications and workload have explored the relationship between the number of Air Traffic Control communication events and their associated workload considering the number and total duration of the communications. However these studies have always been based on the use of a relatively small sample, hence reducing the statistical significance of their results.To overcome this limitation, this study has taken advantage of a highly reliable Automated Speech Recognition system based on the use of semantic analysis for controller event detection developed by CRIDA for AENA (Spanish Air Navigation Service Provider). This system, named VOICE, has enabled the use of a larger data sample (381 hours of recordings and transcripts of Air Traffic Controller Officer voice communications) to analyse the full spectrum of control operational events related to workload. The analysis has focused on two parameters: number of communications and duration of the communications in order to analyse the relationship between the ATC voice communications and the workload in en-route sectors.**



**Keywords-component; ATC voice communications, human factors, workload, automated speech recognition, complexity.**


## I. INTRODUCTION

In the past, several studies have analyzed with different levels of detail the relationship between WL (Workload) (and/or taskload) and ATC (Air Traffic Control) voice communications [1]. The goals of these studies were mainly the prediction of subjective WL [2-4] and the exploration of its relationship with traffic complexity, being of particular relevance for the purpose of this paper is a network dynamics based approach [5]. All of them have in common the absence of the voice communication content, or semantic meaning, plus the common difficulty to use a wide sample of ATC voice communications, either by its unavailability of it or due to the high cost (in terms of time and effort, largely increased if semantic analysis was required) that a proper analysis, segmentation, transcription and categorization would mean.

While some of these problems would (in some cases) have been solved with an ASR (Automated Speech Recognition) system (as usually stated in the conclusions of those studies), traditionally there has not been an efficient system able to do it within the ATC environment, particularly in the operational domain. In particular, several applications have been successful in simulation environment [6, 7], but always relying on the availability of contextual information -such as flight plans-, which is not always accessible. On the other hand, no direct adaptation of the existing ASR COTS (Commercial-off-the-shelf) has been successful, mainly because of the particularities of ATC communications that are later described in this paper.

During the last years AENA has developed an ASR prototype that applies in an innovative way these technologies to the ATC operational domain, in particular from an ATCO (Air Traffic Controller Officer) WL measurement perspective, with focus in the DCB (Demand Capacity Balance) process. The prototype is able to identify and interpret controller-to-pilot and controller-to-controller voice communications in a real operational environment, automatically transcribe voice recordings and determine the associated controller event. This system, named VOICE is currently obtaining high detection rates: (Word Detection Rate, WDR, above 85%) which have allowed its integration with operational ATC communications systems for the automated detection of controller voice events that are later used for post-operational controller WL estimation [8,9]. Thus, this is prototype is able to provide a large data set of interpreted voice communications in an automated way that can be later used for extensive analysis.

WL is widely recognized as one of the key factors affecting controller's performance and thereby system capacity. However, as WL measurements cannot be performed through direct means, they have to be inferred or estimated. Many methods have been developed during the years, covering from a very early extended network queuing analysis [10] and controller's physiological variables monitoring [11] to more recent indicators of traffic complexity [12-15] that are based on the cognitive WL as identified in Wickens studies [16].

It is important to highlight that this prototype works in a non-contextual mode, without external information about the actual traffic. Therefore it allows a reliable ASR in the ATC domain with a high WDR. Non contextual mode permits voice recognition without additional data integration. Applications can use the prototype in a decoupled way such as the one addressed in this study, thereby acting as an enabler for a range of applications using ATC Speech Recognition.

Voice communications can be later correlated with information extracted from post-operational logs of the Spanish ATC Platform resulting in a complete set of controller events usable for WL estimation. Controller events could be detected redundantly either from the interaction with the ATC platform or through voice communications, being remarkable that several events are only obtainable by voice analysis. When the controller event is detected through both sources of information, a cross check is done between them, showing that for some type of events the detection through speech recognition gives more accurate and reliable results. WL is then measured automatically by a prototype named ATON.

As it has already been mentioned, the two parameters used in this study are Average Communication Duration (ACD) and Number of Communications (NoC); in future studies the voice spectrum will also be analyzed. ACD was measured and chosen as an indicator to demonstrate the correlation between WL and voice duration communications.

This paper addresses (i) the methodology used for extracting the average duration and number of communications, (ii) the presentation of the results obtained and (iii) the conclusions and applications of this study.

## II. METHODOLOGY

### A. Rationale

The objective of this study is to analyze the relationship between the ATCO voice communications (ACD and NoC) and WL in order to provide their mathematical relations at different WL levels, providing behavioural patterns that can be used for later modelling.

The key factor of this analysis is a large sample of voice recordings available thanks to the ASR model, made only and exclusively for the ATC domain, allows the production of a large number of automatic transcriptions. To avoid human interference, recordings are generated and collected automatically.

### B. ATC Voice Communication Characteristics

In the ATC domain, before the upcoming of digital communications, voice communication via radio was the primary mean for traffic control. In fact, nowadays it is still the only communication method between controllers and pilots in most control centers.

Communication activity in this paper is defined *as the act in which the ATCO sends a transmission either to an aircraft in the sector under his/her responsibility or to another controller in a different sector*. Empty transmissions are not considered as a communication activity in this model, as they do not generate controller events or WL.

While a large set of different communication channels are included in the CWP (Controller Working Position), the ones affecting controller WL are those which contain input and output controller voice communications: the controller headset output channel PJ (Panel Jack), RD (Radio Channel) being an input line through which the controller receives the aircraft communications in the sector, and coordination TF (Telephone Line), a bidirectional channel specifically used to coordinate control actions between adjacent sectors/controllers. The first one, a purely output channel, contains the highest percentage of communications originated by the controller either to pilots or rarely to other controllers; whereas the coordination line is occasionally used and always to communicate between controllers.

A key aspect of the ATC communications analyzed in this case of study is that they could be either in Spanish, English, or even both mixed in the same sentence. This particular feature due to the fact that both Spanish and English are official ICAO (International Civil Aviation Organization) languages. In practice, this implies that the speech recognizer must be able to detect communications in any of the two languages without having previous notice of which one of them is actually being spoken.

### C. Data Used

The data used in this study has been obtained from the Madrid ACC (Area Control Centre) applied to operational voice recordings. To respect confidentiality, a full year of recordings was used without identifying the initial and final dates. The number of raw hours in the sample is 862 hours, of which around 44% of them contain voice communications (381 hours), meaning a significant set of data, fully sufficient in terms of completeness for the purpose of the study.

A wide set of different sectors (all en-route) have been considered for these communications, with samples obtained at different times of the day from different days of the year with different traffic amounts, thereby covering a fully representative sample adequate for a representative and significant analysis.

### D. System Architecture

ATON is a prototype that processes information from both the Spanish ATC Platform (post-operational) and voice communications to obtain accurate information for statistical computing and automated data and metrics.

ATON has access to the Sectorization Data, Flight Plan Data and Radar Data for Madrid ACC twenty four hours after the operation time. However, in some situations the analysis of this data (such as the sector entries) is insufficient and needs to be complemented with ATC voice communication. The right combination of both sources provides the information needed to extract all the control events.

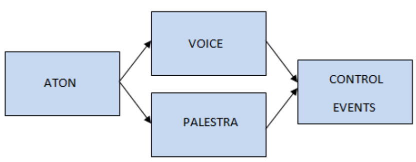



Figure 1 : System workflow

The VOICE module of ATON, after several iterations and intensive training, is provided with acoustic and language models tailored specifically for the Madrid ACC. These models have a word detection rate of approximately 86%. For the automated and secure voice communications recording, the system chosen is NICE[1]. Voice information is encoded in ADPCM32[2] as it has a recognition rate similar to the uncompressed signal and consequently a substantial improvement in the quality is obtained. As stated, for each UC, there are three channels; the one of interest for the study is the PJ line, which contains controller-to-pilot communications.

### E. *Workflow*

In search of a relationship between voice communications and controller WL, an analysis has been performed with all the voice communications of a year in terms of duration and frequency for relevant sets of WL depending on the sector evolution index: sectors with bigger WL measures show a small evolution index. According to AENA WL calculation methodology, sector are categorised according a metric called evolution index, defined as the average of the additional time in evolution percentage ($\%T_{evol}$), the action in evolution percentage ($\%Act_{evol}$) and the complexity percentage in evolution ($\%Comp_{evol}$):

$$i_{evolution} = \frac{\%T_{evol} + \%Act_{evol} + \%Comp_{evol}}{3}$$

For en-route sectors, depending on this index, three sector classifications can be done:

1) *Group 6: High Evolution Sectors* $i_{evolution} > 50$
2) *Group 7: Low Evolution Sectors:* $50 > i_{evolution} > 15$
3) *Group 8*: *Route:* $i_{evolution} < 15$

Groups 1-5 correspond to Terminal Maneuvering Area (TMA) and are out this study. Breakpoints for the grouping are obtained through empirical observation and are obtained from AENA WL calculation methodology.

The audio files obtained are processed within the VOICE server producing an XML output file. Thus the need to store or remove the voice recordings from the centre is eliminated. The XML output files (text files) are sent encrypted to ATON. The information of interest to this study is: the *control event*, the *UC*, *date* and *time*, and *duration*. Each XML output files usually last for an hour (normally o'clock to o'clock).

The first step of the procedure aims to univocally identify all the voice control events in one hour; it then associates them with their related WL estimation as well as the number of controlled flights, thus getting rid of the "null[3]" events. Of a total of 281,506 available events, 25.92% were null events.

[1] NICE: Recording and storage voice system.

[2] ADPCM32: Adaptive differential pulse-code modulation (ADPCM) is a variant of differential pulse-code modulation (DPCM) that varies the size of the quantization step, to allow further reduction of the required bandwidth for a given signal-to-noise ratio. For 32 kbits per second, corresponds to 4 bits.

[3] "Null" events do not have an associated event because the event has not been recognized or because theses phrases do not have a control event inside.

The second step is to calculate the duration average and the duration standard deviation for all the events associated with each WL measure. In this study, 862 WL samples are used with a total of 208,535 not "null" events. General information of the exercise without "null" events is shown in Table I.

TABLE I. INFORMATION OF THE EXERCISE

| | Total |
|---|---|
| **Length of the exercise (s)** | 1,173,969 |
| **Number of the flights identified in the exercise** | 23,698 |
| **Number of communication events in the exercise** | 208,535 |

## III. RESULTS

Datasets corresponding to Groups 1 to 5 have been filtered because they relate to the TMA and this study focus on the en-route flight phase.

WL measurements are evaluated and classified into the three already mentioned groups. A diverse WL sample has been used for this study to cover all the cases. Groups 6 and 7 have smaller WL values as all the events in this kind of sectors are usually the same. In Group 8 the WL values become higher due to the higher WL weights associated to typical events (i.e., headings). Controllers in low evolution sectors have to provide additional effort when re-routing a single flight.

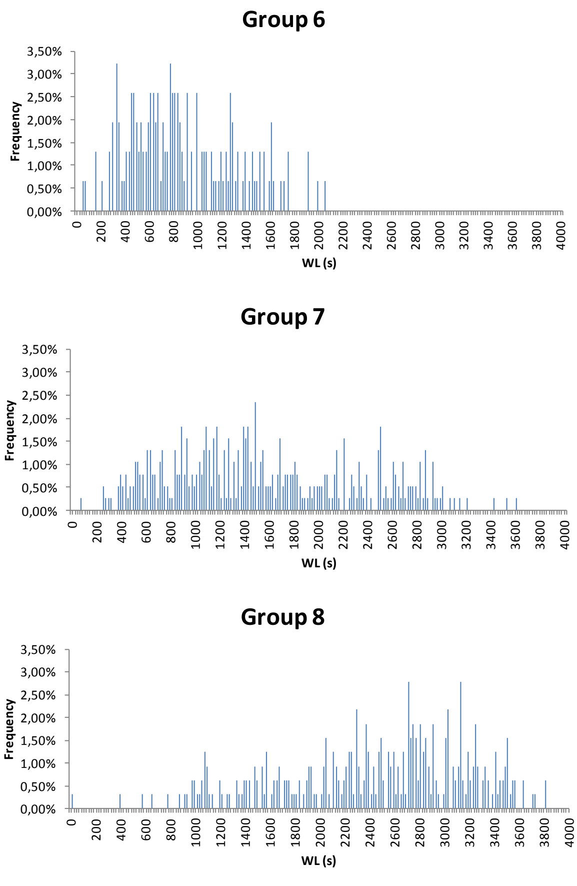



Figure 2. WL Analysis

The results of the data analysis performed are divided into event types that the VOICE "understands[4]". The Event Detection Rate without callsign (EDRno_callsign) results as well as other relevant parameters are displayed in Table II[5]:

TABLE II. DETAILED EVENT DETECTION RESULTS

| Event Code | Event Description | EDRno_callsign | Communication Duration | |
|---|---|---|---|---|
| | | | **$\mu$** | **$\sigma$** |
| Ac12v | SSR Frequency change Communication | 58.3% | 6.03 | 2.65 |
| Ac13.1v | STAR assigment Communication | N/A | N/A | N/A |
| Ac3.4.11v | Clearance or instruction Communication | N/A | N/A | N/A |
| Ac7v | ILS Authorization Communication | 75.0% | 7.03 | 3.00 |
| Ac9v | Essential information Communication | 38.8% | 7.49 | 3.55 |
| Av | Level Change Communication | 93.8% | 5.45 | 2.91 |
| Cov | Inter-sector controller-controller coordination | 17.1% | 11.7 | 8.82 |
| CRv | Clearance/authorization Correction communication | 100% | 11.2 | 2.81 |
| Csv | Sector Change Communication to Pilot | 87.6% | 5.47 | 2.5 |
| CTEv | Sector Entry Communication | 95.1% | 4.96 | 2.46 |
| Dv | Direct Communication | 80.0% | 5.68 | 3.56 |
| H1v | Holding Stack Communication | 33.3% | 7.28 | 3.56 |
| Sv | Heading Type 2 Communication | 100.00% | 6.92 | 3.98 |
| Vv | Speed Change Communication | 20.00% | 6.81 | 3.84 |
| Xv | Heading Type 1 Communication | N/A | N/A | N/A |

The study has focused on the events that should have a bigger impact on the WL (i.e. the most frequent; for this study the break value has been established at 87% detection rate).

The data has been fitted using regression analysis, considering the minimum square fit. Linear, logarithmic, and quadratic functions have been analysed, being in every case the quadratic approach the most accurate. The relation between WL level and voice communications (number and average duration indicators) is then represented as:

$$y = \alpha x^2 + \beta x + \gamma$$

The representation for each event is done considering the 5% and the 95% percentiles in order to avoid distortions in the regression analysis. It is important to remark, that there are some cases in which the WL has also a value, but there are no voice communication recordings. That is the result of situations when the controller is no talking, but he is assuming flights in the screen or solving potential conflicts.

[4] VOICE is currently able to recognise fifteen different control events.

[5] It is considered separately each type of detected controller event only for the en-route flight phase.

TABLE III. $\alpha, \beta$ *and* $\gamma$ VALUES

| Event | Average Duration Quadratic Approach | | | Number of Communications Quadratic Approach | | |
|---|---|---|---|---|---|---|
| | **$\alpha$** | **$\beta$** | **$\gamma$** | **$\alpha$** | **$\beta$** | **$\gamma$** |
| Ac12v | -88.1 | -29.8 | 2,558.3 | -29.3 | 455.7 | 1,008.5 |
| Ac9v | -48.1 | 58.2 | 2,240.2 | -39.6 | 355.2 | 1,399.0 |
| Av | -46.9 | 329.4 | 1,423.2 | -0.8 | 67.2 | 731.8 |
| Csv | 213.9 | -2,498.0 | 8,971.1 | -0.9 | 110.9 | -791.7 |
| CTEv | -50.8 | 203.8 | 2,038.6 | 1.4 | -18.3 | 1,189.5 |
| Sv | -4.0 | -10.1 | 2,213.3 | -29.4 | 200.5 | 1,738.9 |
| Vv | -1.1 | 10.7 | 2,019.1 | -181.9 | 708.8 | 1,289.7 |

The representation for each event is done considering for the lowest values the 0.05 percentile and for the highest ones the 0.95 percentile in order to avoid distortions in the regression analysis. It is important to remark, that there are some cases in which the WL has also a value, but there are no voice communication recordings. That is the result of situations when the controller is no talking, but he is assuming flights in the screen or solving potential conflicts.

**Sector Change Communication to Pilot (Csv)**

This kind of event represents the 34.45% of the sample. In Figure 3(A) it is clearly observed that the average duration of voice communications increases when the controller WL decreases. Figure 3(B) shows that when the NoC is high, the WL has also a big value. The shape of the duration distribution resembles a Poisson distribution.

**Sector Entry Communication (CTEv)**

Entry events are identified only through voice communications because radar flight track does not allow knowing the entrance exact moment. They represent a 25.18% of the sample. In the one hand, Figure 4(A) presents a high negative gradient, similar to Csv, between the WL and the ACD. On the other hand, the NoC increases as the WL and the movements increase, Figure 4(B). The average duration is less than in sector change communications: in this case, controllers do not have to assign a frequency for the subsequent sector.

**Level Change Communication (Av)**

These events represent a 23.85% percentage of the total of the events in the sample. There is a clear relationship between the ACD and NoC with the WL. But in this case, the slopes in Figures 5 (A) and (B) are less steep than for the sector change communication and sector entry communication because in the level change control events, the ATCO has to think in a specific level and see in the screen if there is a possible conflict with other flights. The value for the standard deviation is increasing and the Poisson distribution becomes flat.

**Essential Information Communication (Ac9v)**

This type of events represents just a 2.23% percentage in the data sample, being in this case the duration higher because of the messages complexity (providing information about traffic, weather and unforeseeable events). The number of communications of this kind is low, while still significant. The relationships observed in the other events are applicable here.

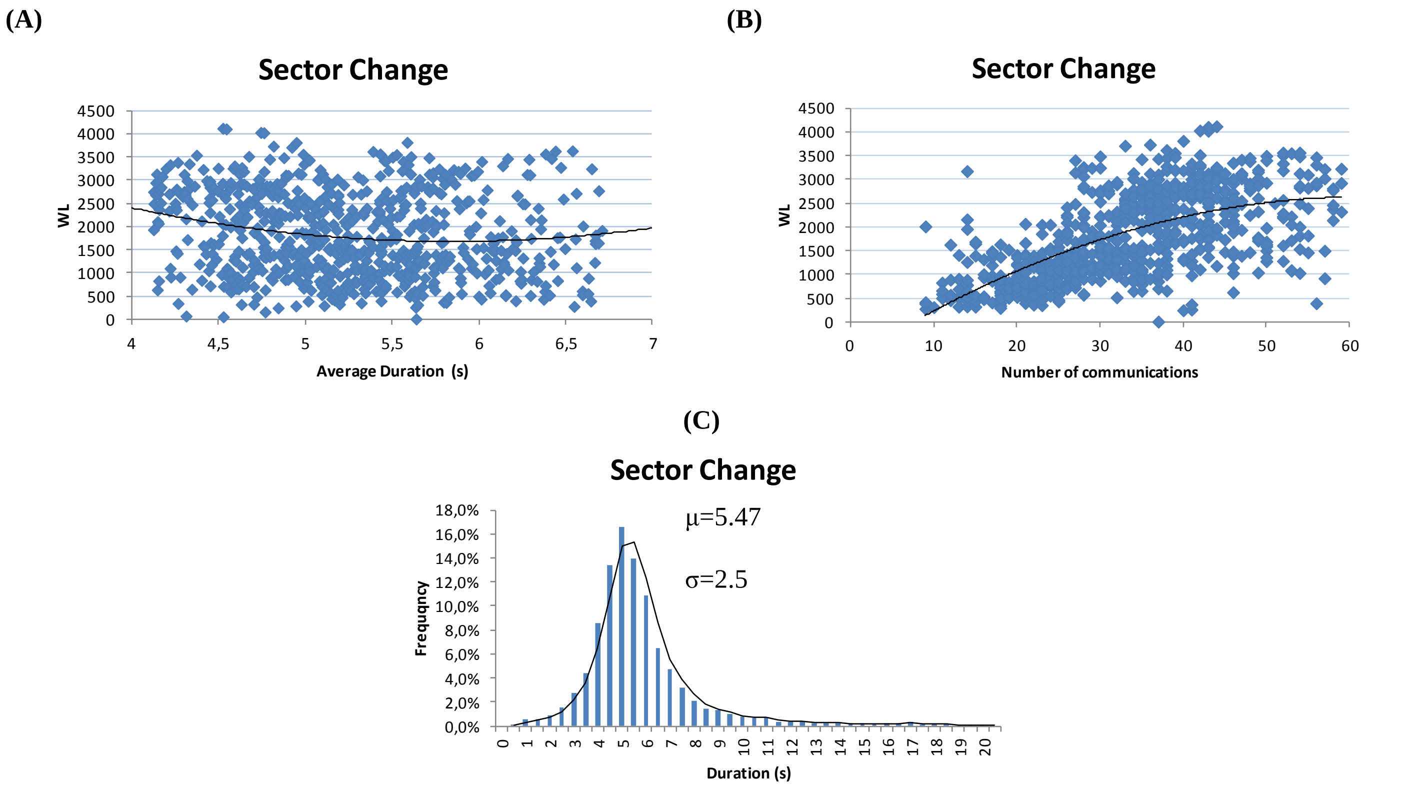



Figure 3. Analysis of all the Sector Change Communication. (A) Relation between WL and average duration (in seconds). (B) Relation between WL and number of communications. (C) Distribution of average duration (in seconds).

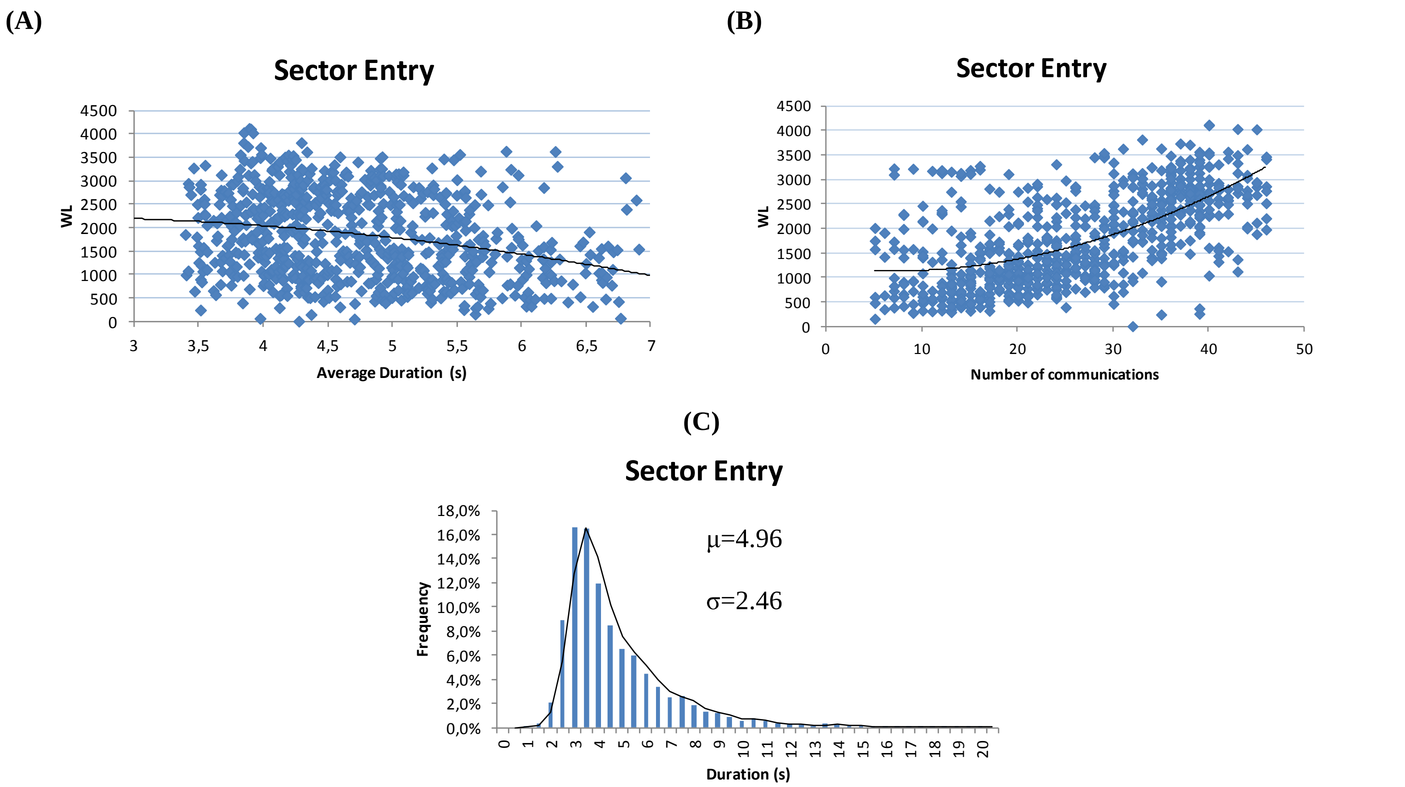



Figure 4. Analysis of all the Sector Entry Communication. (A) Relation between WL and average duration (in seconds). (B) Relation between WL and number of communications. (C) Distribution of average duration (in seconds).

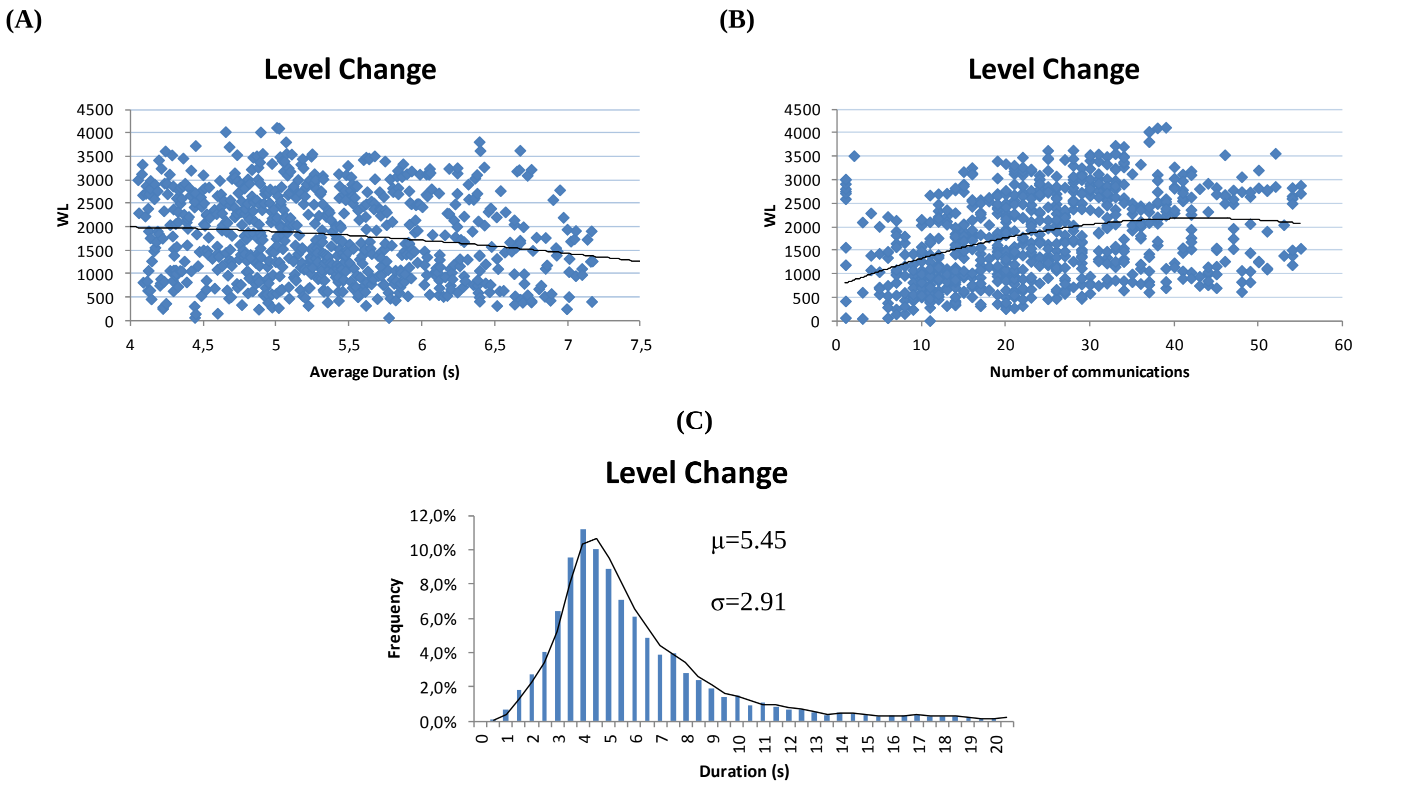



Figure 5. Analysis of all the Level Change Communication. (A) Relation between WL and average duration (in seconds). (B) Relation between WL and number of communications. (C) Distribution of average duration (in seconds).

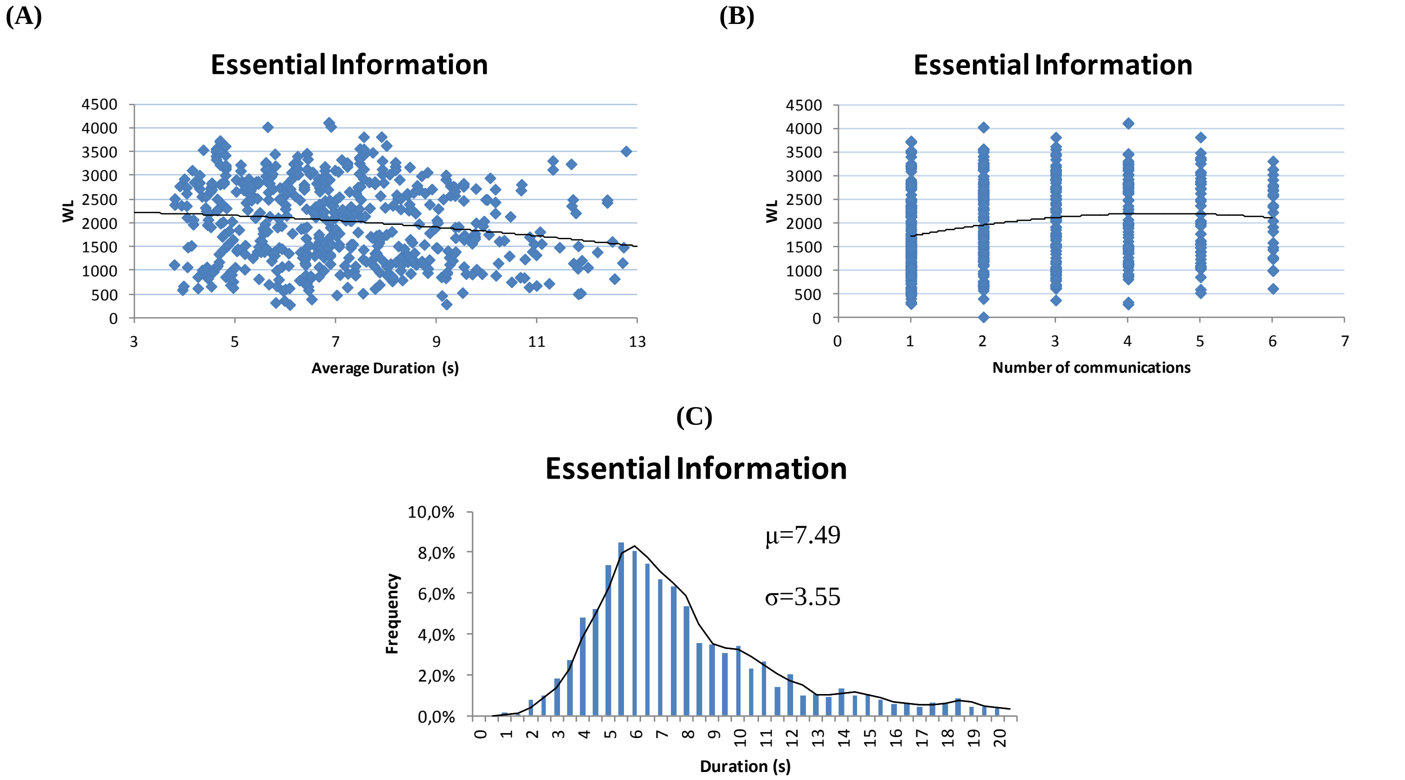



Figure 6. Analysis of all the Essential Information Communication. (A) Relation between WL and average duration (in seconds). (B) Relation between WL and number of communications. (C) Distribution of average duration (in seconds).

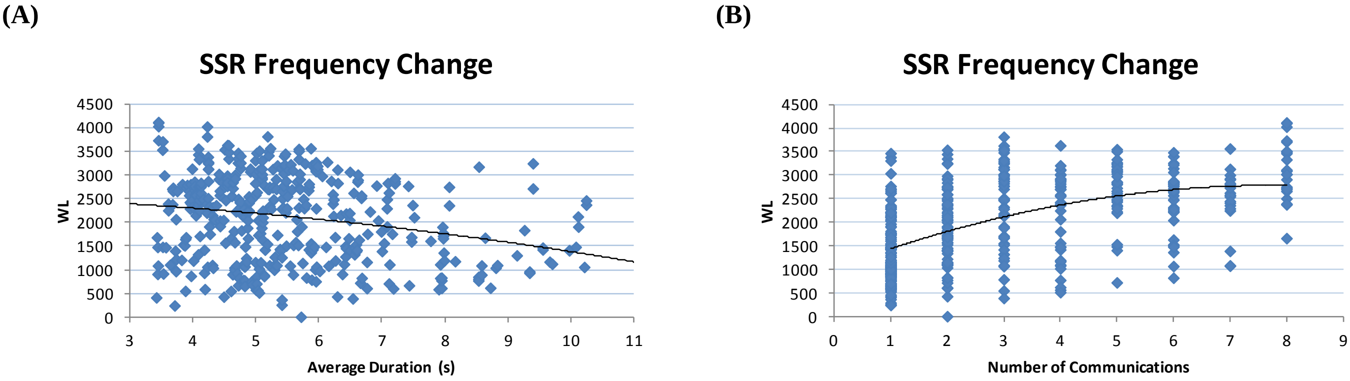



Figure 7. Analysis of all the SSR Frequency Change Communication. (A) Relation between WL and average duration (in seconds). (B) Relation between WL and number of communications. (C) Distribution of average duration (in seconds).

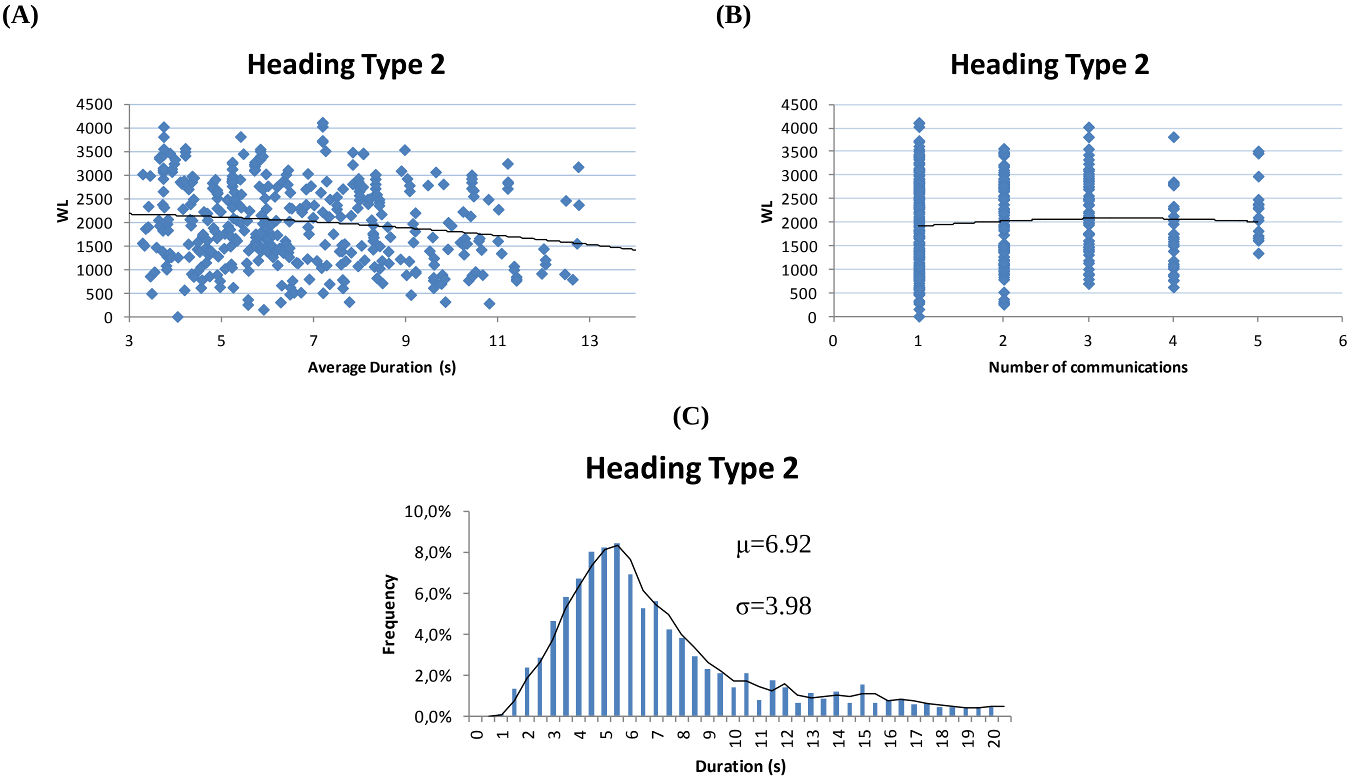



Figure 8. Analysis of all the Heading Type 2 Communication. (A) Relation between WL and average duration (in seconds). (B) Relation between WL and number of communications. (C) Distribution of average duration (in seconds).

### SSR Frequency Change Communication (Ac12v)

For this event, a noticeable relationship between the WL and the ACD and NoC has been observed, as stated before. This kind of events, however, represents only a 1.74% percentage. This (as well as other) unfrequent events display a WL-NoC graph centered in the number of communications, thus emphasizing the discreteness of this axis.

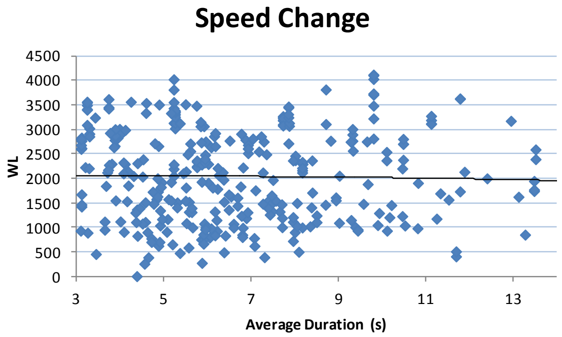
### Heading Type 2 Communication (Sv)

They represent 1.29% of controller events in the sample. For this type of event there is a slight relation between the WL and the ACD, not evident. The curve slope is not as high as in the previous ones. This event associated communication requires the number of degrees for the heading, what implies its average duration to be larger than in other similar events..

### Speed Change Communication (Vv)

They represent 0.76% of the sample and the dependency of the workload and the two variables defined, if exists, is really loose.

**(A)**




**(B)**

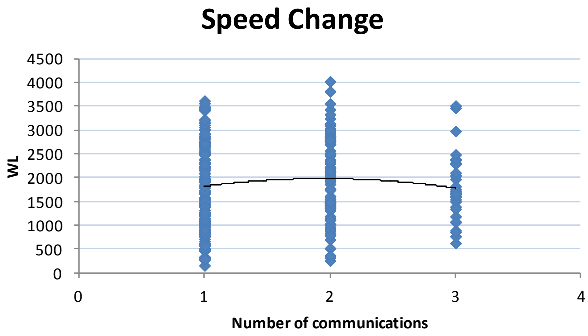



Figure 9. Analysis of all the speed Change Communication. (A) Relation between WL and average duration (in seconds). (B) Relation between WL and number of communications. (C) Distribution of average duration (in seconds).

## IV. CONCLUSIONS

This paper explores the relationship between Air Traffic Controller Officer Voice communications and WL through a wide and representative set of real operation ATC communications, showing that a relationship exists between both magnitudes.

Linear, logarithmic and quadratic approached have been analysis to determine the most accurate approach, being the last one the best option in every case. However, this relationship is not homogeneous amongst events. Events identified as "Sector Change Communication to Pilot", "Sector Entry Communication", and "Level Change Communication" show the steepest change in WL / voice communication duration. In the linear approach, the steep is a good indicator of the strength of this relationship.

WL estimations can be improved if the voice related impacts are taken into account within the WL cognitive model. Parameters such as communication duration based on the identified event type, as well as the frequency of the communications will change significantly the estimation of the cognitive effort associated to the WL.

The use of structured communications, in which there is little or none deviation from a standard template result in smaller communication durations and thus in smaller WL intervals. This signals a possible application related to the automation of some of these tasks using Controller Pilot Data Link applications.

A further stage of the study will identify and elaborate a more sophisticated voice parameters-WL model that includes more complex non-linear relationships and to define indicators for the voice spectrum.

## ACKNOWLEDGMENT

The authors wish to thank AENA and CRIDA for providing all the data used in this study.

## REFERENCES

[1] D.H. Poterfield, "Evaluating controller communication time as a measure of WL", International Journal of Aviation Psychology, pp. 171-182, 1997

[2] C. Manning, and E. Pfleiderer, "Relationships between measures of air traffic controller voice communications, taskload, and traffic complexity", 5th USA/Europe Air Traffic Management R&D Seminar, 2003.

[3] C. Manning, S. Mills, C. Fox, E. Pfleiderer, and H. Mogilka, "The relationship between air traffic control communication events and measures of controller taskload and WL", 4th USA/Europe Air Traffic Management R&D Seminar, Santa Fe, 2001.

[4] K. Cardosi, "Time required for transmission of time-critical air traffic control messages in an en-route environment", International Journal of Aviation Psychology, vol.7, pp.171-182, 1993.

[5] Y. Wang, M. Hu, P. Bellot, F. Vormer, and V. Duong, "Spatial, temporal, and grouping behaviors in controller communication activities", 9th USA/Europe Air Traffic Management R&D Seminar, Berlin, 2011.

[6] G. Churcher, "Speech Recognition for Air Traffic Control", Technical Report, Leeds University, School of Computer Studies, 1996.

[7] D. Schäfer, "Context-Sensitive Speech Recognition in the Air Traffic Control Simulation", 4th USA/Europe Air Traffic Management R&D Seminar, Santa Fe, 2001.

[8] J.M. Cordero, M. Dorado and J.M. de Pablo, "Automated speech recognition in ATC environment", ATACCS' 2012, London, 2012.

[9] J.M. Cordero, N. Rodríguez, M. Dorado and J.M. de Pablo, "Automated Speech Recognition in Controller Communications applied to WL Measurement", 3rd SESAR Innovation Days, Stockholm 2013.

[10] D.K. Schmidt, "A queuing analysis of the air traffic controller's WL", IEEE Transactions on Systems Man and Cybernetics, vol. SMC-8, no. 6, pp. 492-498, 1978.

[11] P. Averty, S. Athènes, C. Collet and A. Dittmar, "Evaluating a new index of mental WL in real ATC situation using psychophysiological measures", 21st Digital Avionics Systems Conference, 2002.

[12] B. Sridhar, K.S. Sheth, and S. Grabbe, "Airspace complexity and its application in air traffic management", 2nd USA/Europe Air Traffic Management R&D Seminar, 1998,

[13] C. Manning, and E. Pfleiderer, "Relationship of sector activity and sector complexity to air traffic controller taskload", FAA Oklahoma Technical Report, 2006.

[14] M. Cano, P. Sánchez-Escalonilla, M. Dorado, "Complexity Analysis in the next generation of Air Traffic management system", 26th Digital Avionics System Conference, 2007.

[15] S. Athènes, P. Averty, S. Puechmorel, D. Delahaye, and C. Collet, "ATC complexity and controller WL: trying to bridge the gap", HCI-Aero, 2002.

[16] C.D. Wickens, Engineering Psychology and Human Performance, Harper-Collins, 1992.

[17] Eurocontrol, "Evaluation of the human voice for indications of workload induced stress in the aviation environment", December 2006.