THE COMPUTATION OF COMMUNICATION COMPLEXITY IN AIR TRAFFIC CONTROL MESSAGES

O. Veronika Prinzo, FAA Civil Aerospace Medical Institute, Oklahoma City, OK

Al Hendrix and Ruby Hendrix, Hendrix Consulting Services, Roswell, NM

Abstract

Over 10 years have passed since a comprehensive analysis quantified the types and frequency of readback errors and communication problems that occur in the air traffic control (ATC) environment. Hence, a content analysis was performed on approximately 50 hrs of pilot and controller messages that were transmitted from five of the busiest US terminal radar approach control (TRACON) facilities between Oct 2003 and Feb 2004. This report describes the computation of ATC message complexity. Furthermore the effects of ATC message complexity and message length on pilot readback performance are presented. The findings show that pilots experienced more difficulty reading back ATC high complexity messages while performing approach tasks as compared with departure tasks. The effects of message length on readback performance were apparent only for approach tasks. Also, nonstandard phraseology associated with a lack of English language proficiency and international communications were present. With increases in international travel, areas of concern related to English language proficiency and language production need to be addressed.

Introduction

The success of Air Traffic Management (ATM) rests upon effective communication in the development, construction, delivery, deployment, and evaluation of its various systems and processes. The operational communications between controllers and pilots are a critical component that should not be overlooked or taken for granted. The phase of flight is relatively unimportant — it could be that the pilots are attempting to communicate with controllers at airports, or in terminal, en-route, or oceanic airspace. What is important is that communication occurs.

The accurate transmittal and receipt of information is critical for communication to occur. The source and its destination may or may not be human. For example, a primary radar transmitter uses a continually rotating tower-mounted antenna to broadcast electromagnetic waves (pulses) of which a proportion is reflected back from objects that are within range of the transmitter. The antenna's orientation provides the object's bearing while the time taken for the pulse to reach the object and return measures its distance from the transmitter. For aircraft, bearing and distance information are converted into ground position information that is presented to the controller on a situation display as moving flashes of light. Nothing is required of the aircraft to provide information back to the radar.

The act of transmitting information from source to destination is an important step in communicating. However, the echoing back of source information is not communication. A critical element of communication requires a common ground of understanding complete with the acknowledgment that the information was received and understood.

Pilot-controller voice communication is built upon a redundant readback-hearback loop. To ensure accurate communication, controllers send instructions and clearances to pilots who read them back. During readback, controllers listen for accurate reproductions of their original messages. The presence of a mistake is a readback error. Occasionally, pilots request complete or partial repeats of original messages, request clarification, or indicate that they cannot comply. Controllers and pilots actively participate in the exchange of information, ensure its accuracy and are committed to providing safe flights.

Readback errors are more likely to occur during approach when pilots experience the most challenging aspects of their flights and ATC messages are more complex and lengthy [1]. The failure to communicate between captains, first officers, flight crew, and controllers makes itself known when quality assurance personnel listen to and evaluate voice tapes to determine the causal or contributing factors of operational errors, pilot deviations, runway incursions, mishaps or accidents.

It is common knowledge that humans have limitations in the amount of information that they can successfully process, store, recognize and recall. At first a person forms many groups or 'chunks' with few bits of information per chunk. With learning and experience, more bits are included per chunk. The

upper limit human verbal working memory (VWM) can successfully recode and retain is between five to seven chunks at a time before forgetting occurs.

Consider the ability to translate sound into meaning. Gradually, we learn to organize sound into progressively larger groups by translating them into a verbal code [2]. The narrative below illustrates the concept of recoding into progressively larger chunks:

"A man just beginning to learn radio-telegraphic code hears each *dit* and *dah* as a separate chunk. Soon he is able to organize these sounds into letters and then he can deal with the letters as chunks. Then the letters organize themselves as words, which are still larger chunks, and he begins to hear whole phrases. I do not mean that each step is a discrete process, or that plateaus must appear in his learning curve, for surely the levels of organization are achieved at different rates and overlap each other during the learning process. I am simply pointing to the obvious fact that the dits and dahs are organized by learning into patterns and that as these larger chunks emerge the amount of message that the operator can remember increases correspondingly. In the terms I am proposing to use, the operator learns to increase the bits per chunk."

At the onset of an utterance, the sounds at the beginning enter into a limited-capacity VWM where they are processed and temporally stored as phonological representations. That is, acoustically relevant sounds are extracted and encoded into phonemes (i.e., consonant-vowel-consonant clusters) that form syllables (e.g., stress patterns and intonation) that are assembled to create words, phrases, clauses, and other constituents. These representations must be maintained in an active state (rehearsed) otherwise they begin to decay in about 2 s [3] or are overwritten by incoming information. Furthermore, [3] proposed a linear relationship between the number of words correctly recalled and speech rate. Using mathematical modeling, [4] found that 95% of the variance in memory span[1] performance for words, digits, and colors was related to the number of items that were spoken in 2 s.

The phonological-loop model of VWM [5] has demonstrated that the ability to recall accurately information in the order in which it was originally heard is better for word sequences that have shorter as compared with longer articulatory durations (i.e., the time needed to pronounce the word sequence). This effect holds true when two sets of words are matched in the numbers of phonemes and syllables but differ in mean articulatory durations [6, 7]

---

[1] Memory span refers to the number of items (usually words or digits) that a person can hold in working memory. Tests of memory span are often used to measure working memory capacity. The average span for normal adults is 7.

An utterance's complexity can be derived from its grammatical weight — the amount of information expressed in its constituents as measured by the number of words, syntactic nodes, or phrasal nodes in the constituent [8]. To successfully recode sensory information into progressively larger chunks requires automatic recoding [2]; otherwise, as new inputs are being transmitted, they will be sacrificed while attempting to retain the name of the last group.

Communication elements are the fundamental units of meaningful verbal language [9]. Within aviation, they are identified by their functionality; i.e., their purpose, operation, or action (Address/ Addressee, Advisory/Remark, Request, Courtesy, Instruction/Clearance, and Non-Codable) and are restricted with regard to their aviation topic (AT, e.g., altitude, heading, speed, traffic, route, etc) [10].

We refined the computation of level of complexity [11] in ATC communication elements. Furthermore, we define the level of complexity of a communication element as each word or set of words that contains a new piece of information critical to the understanding of that communication element. As is often the case, ATC messages contain multiple communication elements and message complexity is the sum of the values assigned to each message.

As noted in [12], ATC prescribes that controllers use a rigid set of words/phrases. This phraseology narrows the definition and meaning of communication elements. Some of these words and phrases serve as anchors that make the communication element more precise in its interpretation. Some anchors attach meaning to the numbers. For example, the significance of "3-5-0" is ambiguous until an anchor word appears with it — "3-5-0" can easily be interpreted as a heading, altitude, or speed. Thus degrees are associated with heading, knots with speed and descend/climb with altitude. When so used, anchors assist in the interpretation of communication elements and restrict the meaning assigned to ATs.

Our scoring scheme assigns a complexity value = 1 to each anchor, numerical value, orientation (left, right, center), and the name of fix, point, intersection, marker, etc. as determined by the phraseology usage by the controller according to the examples provided in [12]. Furthermore, it attempts to reflect the added complexity imposed by communication elements that contain more information by assigning them larger values. This assumption holds particularly for altitude instructions. For example, altitude instructions such as 'three thousand five hundred,' 'one-zero thousand' and 'four thousand' most likely impose <u>quantitatively</u>

different loads on working memory. In particular, 'three thousand five hundred' takes longer to pronounce and contains more words than 'four thousand' (e.g., articulatory loop [13]) and utilizes more capacity [2]. When serial reproduction is required, numerical content that utilizes more resources may be partially or completely omitted or lead pilots to request a repetition [14].

Table 1 presents excerpts from the Instruction Complexity Guide [1]. It was developed to increase the reliability and consistency of tabulating complexity for typical ATC phraseology usage. Column 1 presents the AT and Column 2 presents the complexity value (CV). The smaller the CV is, the less complex the phrase. Column 3 presents the phraseology extracted from [12] to support the delivery of that service. In several cases, the phraseology used by the speaker did not appear in [12] (e.g., tight turn, go fast) but was used so frequently, that it was assigned a numerical value.

Anchors appear capitalized, are fixed in their meaning, and designate the to-be-performed action. Italicized words (in parenthesis) are qualifiers that vary with geographical location and aircraft position.

To determine the CV, anchors, qualifiers, and excessive verbiage are assigned a value indicative of new information. In most cases each anchor is counted as one element of complexity. There are several exceptions, however. Some communication elements contain multiple anchors, as is the case 'turn left/right heading (degrees).' The anchor 'TURN LEFT/RIGHT' provides the direction of the turn while 'HEADING' indicates the aircraft's bearing.

Also, qualifiers such as the numbers that comprise an altitude must be evaluated according to the phraseology used by the speaker. That is, the number 'three thousand five hundred' was assigned a CV = 4 (a CV = 1 for each number and a CV= 1 for each anchor) since it would be more demanding than either one-zero thousand (CV = 3) or four thousand (CV = 2). Finally, CV = 1 was assigned to communication elements containing excessive verbiage. Excessive verbiage is determined by comparing the utterance against the phraseology designated in *FAA Order 7110.65P Air Traffic Control*]. If a pilot attempted a verbatim readback, then the coding procedures were applied that evaluated the controllers' transmissions.

An example provided by [11] illustrates how we compute level of complexity. The controller says, "Aircraft XX, change runway to two-five left, cross Santa Monica VOR at or above seven thousand, descend and maintain three thousand five hundred."

We suggest that the transmission contained four ATs: an address, an advisory to expect a change in route, an instruction involving an altitude restriction, and an instruction to change altitude. The altitude restriction had a CV = 5 [cross = 1, point = 1, at or above = 1, numerical value = 1, thousand = 1] and altitude CV = 6 [descend and maintain = 2, numerical value = 1, thousand = 1, numerical value = 1, hundred = 1]. The transmission had a CV = 11.

AVIATION TOPIC	COMPLEXITY VALUE	PHRASEOLOGY
Altitude		3=*(altitude) two digits* +THOUSAND 2=*(altitude) one digit* + THOUSAND 3=*(altitude) two digits* + HUNDRED 2=*(altitude) one digit* + HUNDRED 2=*(altitude) two digits* 1=*(altitude) one digit*
	6	DESCEND/CLIMB & MAINTAIN *(altitude)* THOUSAND *(altitude)* HUNDRED three five
	5	DESCEND/CLIMB & MAINTAIN *(altitude)* THOUSAND one zero
	4	DESCEND/CLIMB & MAINTAIN *(altitude)* THOUSAND four
	*4-8	CONTINUE CLIMB/DESCENT TO *(altitude)*
	*4-8	AMEND YOUR ALTITUDE DESCEND/CLIMB AND MAINTAIN *(altitude)*
	*3-7	AMEND YOUR ALTITUDE MAINTAIN *(altitude)*
	*3-8	DESCEND/CLIMB TO *(altitude)*
	*2-6	MAINTAIN *(altitude)*
	*1-2	*(altitude,* omitted "THOUSAND" "HUNDRED")

Table 1. Excerpt from the for Instruction/Clearance Complexity Guide

METHOD

Materials

Audiotapes. In this report 28 hr 13 min 23 s of approach and 23 hr 56 min 32 s of departure communications were provided by the five busiest US TRACON facilities. There were 12-arrival and 11-departure sectors represented. The amount of voice communications varied from as little as 58 min 55 s on one communication sample to as much as 5 hr 13 min 49 s on another.

The traffic was primarily air carrier, with some private jets, and a few general aviation pilots flying the Coastal VFR Corridor. All sectors had some foreign carriers. Each facility representative was instructed that recordings were to reflect communications-intensive periods during peak traffic loads (as determined by that facility). For the outbound push, the sampled recordings represented morning (7:30 am), afternoon (12:30 pm), mid-day (4:30) and evening (5:54 pm) departures and early-morning (8:45 am), mid-morning (11:00 am), afternoon (12:00 pm), mid-day (3:00, 5:00 pm) and evening (7:15 pm) arrivals during the inbound rush.

Pilot Readback Error Classification Guide. As used here, a readback error is defined as an unsuccessful attempt by a pilot to read back correctly the information contained in the communication elements that comprise the original message transmitted by ATC. As seen in Table 2, the column to the left displays the types of readback errors according to a particular type of AT.

Many of the readback error types are common to all ATs. Presented in the right column of Table 2 are examples of each type of readback error according to the AT in which it was embedded. Preceding each example is the original ATC message. For example, ATC might transmit the following message to AAL10, "American Ten turn left heading two one zero." If the pilot reads back either "three one zero" or "six zero," it would be coded as a substitution error since the numbers in the original heading instruction included neither a three nor a six.

Procedure

Message Encoding. The SMEs met on five separate occasions. Several meetings were used to operationally define message complexity and develop the rules and procedures for encoding each message. This was done to limit the subjective determination of what constitutes information complexity for verbal information.

During the encoding process each transmission was first parsed into communication elements, labeled by speech act category and AT using the procedures developed by [10]. Next the Complexity Guide was used to look up and assign the appropriate CV according to the phraseology used by the controller. The elements of complexity were counted for the a) instructions/clearances speech acts that involved heading, heading modifier, altitude, altitude restriction, speed, approach/departure, frequency, route, and transponder ATs, b) advisory speech act that involved traffic and c) the altimeter portion of weather advisories.

Classification of Readback Errors	Examples
Readback Errors Type (Heading) 1 = Substitution of message numbers 2 = Transposition of message numbers 3 = One type of information read back as another type 4 = Incorrect direction of turn 5 = Omission of one or more numbers 6 = Not assigned 7 = Omission of anchor word(s) 8 = Substitution of anchor word(s)	**ATC "AAL Ten turn left heading two one zero"** 1-"three one zero," or "six zero" 2-"turn left heading one two zero" 3-" two one zero knots" 4-"turn right two one zero," 5-"one zero," "zero on the heading" 7-"two one zero" 8-"two hundred and ten degrees"
Readback Errors Type (Altitude) 1 = Substitution of message numbers 2 = Transposition of message numbers 3 = One type of information read back as another type 4 = Not assigned 5 = Omission of number element 6 = Not assigned 7 = Omission of anchor word(s) 8 = Substitution of anchor word(s)	**ATC "AAL Ten climb and maintain one two thousand"** 1-"to one three thousand" 2-"climb two one thousand" 3-"one two zero knots" 5-"two thousand" 7-"twelve" 8-"up to twelve thousand"

Table 2. Readback Error Guide

Inter-coder Agreement. The consistency of data encoding was evaluated as the transcripts for each of the remaining TRACON facilities were encoded. This was achieved by having the first and second author randomly encode the same set of 25 messages (for each facility) and then computing agreement. In each case it exceeded 95%.

Once all the messages were encoded Krippendorff's alpha [15] was performed on 125 different messages to evaluate inter-coder agreement. It is a statistic that is widely applicable wherever two or more methods of processing data are applied to the same set of objects, units of analysis, or items and the question is how much they agree. Treating the ratings as ordinal data produced an $\alpha = .9898$ indicating high inter-rater agreement.

RESULTS

The results presented here examined the prevalence of pilot readback errors and requests for ATC to repeat all or part of a previous transmission as a function of ATC message complexity and message length (as determined by counting the number of ATs in the transmission) — excluding Address/Addressee and Courtesies. They were derived from 11,159 ATC transmissions. Each ATC transmission that met the selection criterion (i.e., it contained an instruction, advisory, or a combination of instruction and advisory speech acts) was paired with the pilot's response to that message. Statistical significance was evaluated with $\alpha \leq .05$.

Each pilot readback was evaluated for accuracy, and the number of errors present was recorded (e.g., a zero indicated no error while a value of 3 indicated 3 errors). There were 723 individual readback errors present in 688 pilot transmissions — approximately 6% of the pilots' readbacks contained a readback error. Pearson correlations revealed that readback errors increased as the complexity, $r(11159)=.196$ and message length (i.e., number of ATs), $r(11159)=.180$ in a controller's message increased. Likewise, albeit to a lesser degree, the number of pilot requests increased with message complexity, $r(11159)=.020$ and message length, $r(11159)=.054$.

Message Complexity

Approximately 10,471 messages resulted in no readback errors — 93.8% of the pilots' readbacks were correct. For the 6.2% faulty pilot readbacks, 654

contained 1 error and another 34 contained 2 or more errors.[2]

ATC messages with complexity values ≥ 10 were more difficult for pilots to read back correctly, as evidenced by the presence of 2 or more errors per readback. In fact, the percentage of readback errors reached double-digit status once the threshold of 10 was crossed. Prior to reaching a complexity value of 10, the percentage of readback errors was fairly stable — ranging from as little as 2.3% to 6.1%. Message complexity values between 11 and 13 resulted in an increase in readback errors from 10.8% to 19.2%, while complexity values > 16 had an error rate that approached 38%.

Each ATC message was classified as either low (≤ 09) or high (≥ 10) complexity. Each pilot transmission had a readback value, and the average of those values was computed for each aircraft. An ATC Sector (Approach, Departure) by Message Complexity (Low, High) Analysis of Variance (ANOVA) conducted on pilot readback performance revealed that pilots produced more errors while in an approach (Mean = .13 SD = .30) compared with a departure (Mean = .04 SD = .15) sector, $[F(1,3700) = 129.00]$. Also, more complex ATC messages had a higher incidence of being read back incorrectly (Mean = .17 SD = .38) than less complex messages (Mean = .04 SD = .12), $[F1,3700) = 154.39]$. However, these statistically significant main effects must be qualified by a statistically significant ATC Sector by Message Complexity interaction, $[F(1,3700) = 97.18]$ that is presented in Figure 1.
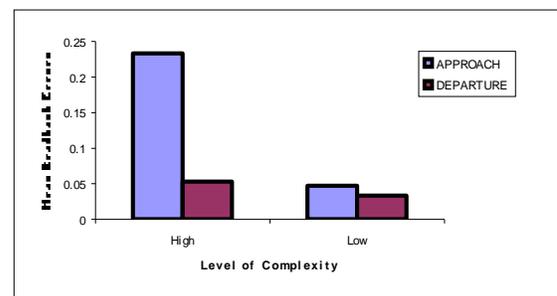

Figure 1. Mean Pilot Readback Errors Presented by ATC Sector and Message Complexity

The Tukey Honestly Significant Difference (HSD) statistic revealed that pilots experienced more difficulty reading back approach control high-

---

[2] Applying a liberal scoring criterion (i.e., partial readback of some numbers in a heading, speed, altitude, or radio frequency and excluding some anchor words such as fixes or points not counted) resulted in 1.3% readback errors.

complexity messages than reading back departure control high-complexity messages or low-complexity messages from either approach or departure control.

Message Length

Messages with one AT occurred for 54.2% of the transmissions, and they resulted in 3.8% readback errors. Messages with 4 ATs appeared in 5.2% of the transmissions, producing 25.7% readback errors. Once again, pilot mean readback performance scores were computed for each aircraft. The results of the ATC Sector (Approach, Departure) by Message Length (1AT, 2AT, 3AT, 4AT) ANOVA revealed that more readback errors occurred when pilots were in the approach (Mean = .11 SD = .31), as compared with the departure (Mean = .034 SD = .16) sectors, $[F(1,5599) = 78.48]$. As expected, the number of readback errors varied with the number of ATs, $[F(3,5599) = 21.62]$. Tukey HSD comparisons revealed that the fewest readback errors occurred when ATC messages contained one AT (Mean = .04 SD = .14). There was no reliable difference between messages with 2 or 3 ATs (2AT = .06 SD = .21; 3AT = .08 SD = .26). However, messages with 4 ATs contained the most readback errors (Mean = .30 SD = .51). These main effects are qualified by a statistically significant ATC Sector by Message Length interaction.

Figure 2 shows that as approach control messages increased from one to between 2 and 3 ATs and 4 ATs that the mean number of readback errors increased accordingly. The effect of message length is apparent only for approach control. There was no discernible difference between readback performance for approach and departure sectors for one AT.
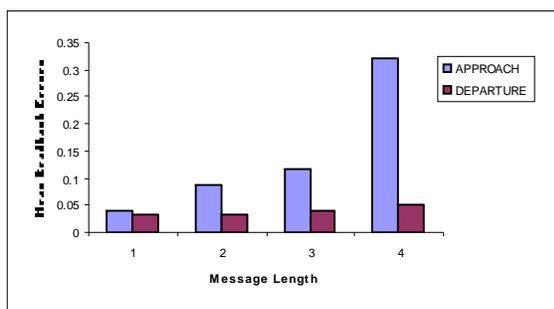

Figure 2. Mean Pilot Readback Errors Presented by ATC Sector and Message Length

Readback Errors and Aviation Topics

Readback errors fall within three major classifications — omission (63.8%), substitution (33.6%), and transposition (2.6%). The distribution of error classes differed across AT. For instance, of the 18.9% anchor word omissions, 12.4% involved heading (e.g., "eight zero"); almost half (11.2% of the 24.6%) of number omissions concerned speed (e.g., "eighty on the speed," "eighty knots"); and over two-thirds of the point/fix omissions related to speed.

Substitution of anchor word(s) and number element(s) represented nearly 75% of the 7 types of substitution errors. Anchor word(s) substitution was more likely to involve altitude restrictions and speed assignments than headings or approach clearances. Similarly, number element(s) substitution was more likely to involve radio frequency, followed by heading and speed instructions. The combination of altitude instructions with altitude restrictions accounted for about 18% of the readback errors involving number element(s) substitution.

Transposition errors involved reordering the number element(s) or point/fix. About 95% involved reversing the order of one point/fix with another.

Hearback Errors

While a pilots' inaccurate readback of a message is called a readback error, a controllers' failure to notify a pilot of a readback error is called a *hearback error*. Of the 723 identified readback errors, approximately 92% were not corrected by ATC.

ATC Corrected Readback Errors. Of the corrected readbacks, 13.8 % involved omission, 79.3% involved substitution, and 6.9% involved transposition errors. It may be that some types of readback errors are more critical than others and may pose a greater risk to safety than others. For example, transposing a number in an AT may be more of a threat in some situations than omitting a number or the substituting an anchor word with its synonym.

A reexamination was performed of the corrected readback errors to compare the opportunity to correct an error with the actual number of corrections made. Only 1.7% of all the omission errors (8/61), 18.8% of the substitution errors (46/243), and 21.1% of the transposition errors (4/19) were corrected.

Coincident Factors to Miscommunications

In this final analysis, transmissions that contained one or more faulty readbacks were examined for the presence of factors that might be correlated with, or have contributed to, its occurrence. Coincident factors included clipped/abbreviated transmissions, nonstandard phraseology, pilot expectation, language barriers, and

transmission overlap (stepped-on, blocked transmissions).

There were 207 pilot readbacks that began with an abbreviated speech act (e.g., "THIRTY HEADING," "EIGHTY SPEED," "ONE ZERO FOUR THOUSAND") that may have resulted from poor microphone technique, poor phraseology, or differences in aircraft radio transceivers. Also, once the pilot began a readback, nonstandard phraseology was another factor associated with 91 transmissions with readback errors. There was a tendency among some pilots to truncate or otherwise abbreviate the numerical values in speed, heading, or altitude assignments. In a similar way, aircraft call signs also were truncated. For example, Ownship67H became Ownship60 and Ownship528 became Ownship520. Some pilots used the "point" designation associated with radio frequencies when reading back altitudes and speeds or substituted "decimal" for the word "point" when reading back a radio frequency. Also, several pilots flying for foreign air carriers displayed some problems in English proficiency and language production. Finally, pilot expectation (n = 16) played a coincidental role in pilot readbacks and was associated with the pilot of one aircraft reading back the contents of a message meant for the pilot of a different aircraft (a 'stolen' transmission).

Pilot Requests for Repeat of Part or all of the Transmission

There are times when pilots are busy setting-up for the approach, completing checklists, or performing other station-keepings tasks, they hear, or think they hear, their aircraft's call sign. Uncertain of the accuracy of an attempted readback, they may request a repeat of all (say again) or part (what was that heading again?) of the message. In other instances, they may request confirmation of the ATs that they thought they heard (confirm we're cleared down to five thousand).

There were 133 messages in the form of a request (45.1%) or confirmation (54.9%). Of the 60 requests made, 18.3% were for a full repeat, 78.4% a partial repeat, and 3.3% asked the controller to identify the recipient of the message (who was that for?). Radio frequency (38%) and heading (17%) assignments were more frequent partial "say agains" than altitude (5%) and route (5%) assignments.

There were 73 requests for confirmation — 4.1% for a full transmission, 65.8% for a specific AT, and 30.1% for the recipient of the message (was that for me?). Of these requests 23.0% were for headings and 16.0% were for altitude assignments.

DISCUSSION

Whether by human or avionics, the accurate transmittal and receipt of information is necessary but not sufficient for communication to occur. Information transfer may be between aircraft using an automatic dependent surveillance-broadcast (ADS-B) to deliver real-time graphically-displayed traffic information to pilots. Or pilots and controllers can use a datalink to send and receive digitized instructions/clearances to each other as text messages. Within these examples is the critical factor that results in communication. Pilots and controllers, the human factors, must acknowledge or otherwise confirm that a common ground of understanding occurred between the source and its intended pilot recipient.

The ideal controller-pilot communication process would show a direct 1:1 relationship between the production of an ATC message and its readback by the pilot-recipient. To avoid the occasion for faulty communication, pilots sometimes jot down the contents of an ATC message on a kneeboard or scratch pad clipped onto the yoke of the aircraft. In commercial aviation, pilots also change the dials on their mode control panel as they receive new headings, altitudes, or speeds; radio frequencies are dialed into their second radio transceiver.

The results presented in this report provide a description and summary of the controller-pilot communication process during normal, day-to-day operations in the TRACON environment. On average, one aircraft requested and received ATC services every 1 min 26 s in the approach sectors and 1 min 6 s in the departure sectors. The number of ground-to-air transmissions averaged 7.3 messages per aircraft for approach control and 4.7 for departure control. Approximately 13 messages were exchanged (from initial contact until the aircraft was switched to the next controller in sequence) that involved an allocation of about 1 min 16 s of *airtime* per aircraft.

As is often the case, ATC messages contain multiple communication elements. The information content in a communication element contributes to its level of complexity. A message's complexity is partially determined by the sum of the CVs assigned to the information content of individual communication elements. Other factors that could affect a message's complexity include message structure, importance, as well as the number of communication elements requiring pilot action.

Some communication elements are ancillary – they do not affect the pilot's ability to aviate or

navigate (e.g., general acknowledgments, greetings). The more important ones provide new information, confirm expectations, verify existing information, or negate that information (e.g., heading, altitude, speed instructions; approach/departure clearances; traffic advisories). For example, [16] reported no apparent correlation between the number of elements and the complexity ranking among a set of 28 ATC clearances on the ability of a sample of airline pilots to accurately copy down previously recorded clearances. They suggest that factors such as familiarity with the operating procedures within a domain (air carrier, general aviation) and geographical location (knowing the names of the navaids, fixes, etc.) affected what pilots in their study copied accurately and what was discarded.

Several investigators have documented the vulnerability of pilot memory and readback performance. For example, [17] reported that message complexity directly affects pilot memory. Several field studies showed fewer pilot readback errors and requests for repeats when controllers' messages were short and simple [18, 19]. Likewise, laboratory studies [14, 20] found that readback errors and pilot requests were more likely to occur in response to longer ATC messages. Finally, the operational data analyzed here provide additional evidence that readback errors and pilot requests increased with increases in complexity and message length (when measured by the number of ATs in a controller's message). Of particular interest (but no surprise) was the finding that pilots experienced the most difficulty reading back ATC messages during the approach segment of their flight. Adding to their workload was the read back of a message with more than one AT or a complexity value ≥ 10 as evidenced by increased readback errors.

Readback errors fell into three major groupings: omission, substitution, and transposition errors. The type of readback error produced seemed to be related to the type of information read back. For example, pilots were more likely to omit an anchor word or phrase when reading back a heading and either excluded a number or left out the point/fix in a speed instruction. They were more likely to substitute an anchor word(s) when reading back either an altitude restriction or speed assignment than a heading or approach clearance. When instructed to switch frequencies, change to a new heading, or alter the aircraft's speed, pilots were likely to substitute numbers. Finally, a majority of the transposition errors involved reversing the order of one point/fix with that of another within the same message.

It was surprising that controllers corrected only 8% of the readback errors. Why were so few corrected? It would seem that during the hearback process, they might evaluate the intrinsic safety component of each readback and then decide whether or not to correct a detected error. It would follow that some communication elements may have little or no impact on safety, and if corrected, add to radio frequency congestion and task load. In such a situation, pilots might not be alerted to their readback errors since controllers monitor aircraft track and position information on their situation displays. In fact, when listening in on a frequency while observing controllers, it is common to hear a controller whisper "close enough" when some readback errors occur. Apparently, such readback errors were not sufficient to warrant another transmission. Consequently, actively correcting a faulty readback might be a conservative process reserved for transmissions having a direct or immediate effect on safety, aircraft performance, traffic flow, or similar factors.

It may be that some types of readback errors are more safety-critical than others — especially when situational factors are taken into account (e.g., reading back "runway four-left approach" when "four-right" was given following the instruction "turn left"). Controllers were more likely to correct transposition errors than either substitutions or errors of omission. By correcting the pilot as soon as possible, the controller can prevent down-stream consequences — such as potential increases in workload and frequency congestion. For example, if radio frequency number substitution errors went uncorrected, the pilot might switch to the wrong frequency. Typically, the pilot comes back on frequency and requests the radio frequency again; the controller gives it, and the pilot reads it back. This adds to the workload and frequency congestion.

Finally, controllers may be less likely to correct pilots' errors of omission than substitution errors since immediacy of reply and context mitigate the potential for misunderstanding created by missing digits ("one seven zero knots" read back as "seventy knots"), anchor words ("one seventy" in response to a speed instruction), or other omissions. Also, prior knowledge (i.e., aircraft slow down on approach and speed up on departure; and at certain speeds fall out of the sky), coupled with redundant visual information (observing aircraft trajectories on their situation displays), assist controllers as they monitor and verify pilot compliance with their instructions.

If controllers actually missed detecting a readback error and safety was compromised, an investigation would occur. Since these tapes were not pulled for study, it is safe to conclude no operational error or pilot deviation ensued from the faulty hearbacks contained in this report. Still, nearly 30% of operational errors result from readback/hearback errors and failures in comprehending instructions [22]. Future communications systems developers should exploit technology to provide controllers with an error-trapping and altering capability (e.g., using speech technology or data link).

The *Aeronautical Information Manual* provides pilots with good information about basic communication techniques, communication procedures, and phraseology. The key concept is that good communication skills promote safety through a mutual understanding between the pilot and air traffic service personnel.

A new trend occurring in ATC communication is the tendency to round the numbers in the call sign and ATs. Some pilots truncated or otherwise abbreviated the numerical values in speed ("TWENTY FIVE KNOTS"), heading ("ONE FOUR" for a heading of one four zero), or altitude assignments ("DOWN TO FIVE HUNDRED"). It is possible that some abbreviations were due to delivery technique or equipment use, while others may reflect a heightened workload [23].

Other forms of nonstandard phraseology were associated with readback errors. It may be that some of the phraseology used (or heard) by pilots during international flights is making its way into the US. Some pilots used the "point" designation associated with radio frequencies when reading back altitudes (e.g., "THREE POINT FIVE" instead of "THREE THOUSAND FIVE HUNDRED") and speeds (e.g., "TWO POINT SEVEN ON THE SPEED" for "TWO HUNDRED AND SEVENTY KNOTS") or substituted "decimal" for the word "point" when reading back a radio frequency. Several pilots flying for non-US air carriers displayed some problems in English proficiency and language production. For example, reading back a speed instruction as "TWO ZERO HUNDRED" instead of "two hundred knots," or responding to "maintain visual from traffic" as "MAINTAIN VISUAL APPROACH."

Conclusions

Safety is the primary objective of the communication phraseology developed for ATC. Much work remains to be done to bring about one standard and understood phraseology. For example,

some controllers in some countries believe that a "Cleared Direct" instruction means that the pilot is expected to fly the currently filed track over the named waypoints "directly" to the airfield. Other controllers in the same country expect a pilot receiving a "Cleared Direct" instruction to deviate from the previous route clearance and fly on a straight track between the aircraft's present position and the point mentioned to which the aircraft has been cleared. Also, phrases such as "Line up and wait" and "Taxi into position and hold" create problems. With increased international travel and the gradual migration of other phraseologies into its vocabulary, pilots and controllers must remain vigilant in the accurate production and recitation of ATC messages.

References

[1] Prinzo, O.V., A. Hendrix, and R. Hendrix, 2006, *The outcome of message complexity on pilot readback performance,* Report no. DOT/FAA/AM-06/25, Washington, DC: FAA.

[2] Miller, G., 1956, The magical number seven plus or minus two, Some limits on our capacity for processing information, *Psychological Review*, *63*, 81-97.

[3] Baddeley, A., N. Thomson, and M. Buchanan, 1975, Word length and the structure of short-term memory, *Journal of Verbal Learning and Verbal Behavior, 14,* 575-589.

[4] Schweicker, R. and B. Boruff, 1986, Short-term memory capacity: Magic number or magic spell? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 12,* 419-25.

[5] Baddeley, A., 1987, *Working Memory,* Oxford, England: Oxford University Press.

[6] Baddeley, A. and G. Hitch, 1974, Working memory, In G. Bower (Ed.), *The psychology of learning and motivation* (Volume 8, pp. 47-89), New York: Academic Press.

[7] Mueller, S., T. Seymour, D. Kieras, and D. Meyer, 2003, Theoretical implications of articulatory duration, phonological similarity, and phonological complexity in verbal working

memory, *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29,* 1353-80.

[8] Wasow, T., 1997, Remarks on grammatical weight, *Language Variation and Change*, *9*, 81-105.

[9] Prinzo, O., (1996), *An analysis of approach control/pilot voice communications,* Report no. DOT/FAA/AM-96/26, Washington, DC: FAA.

[10] Prinzo, O., T. Britton, and A. Hendrix, 1995, *Development of a coding form for approach control/pilot voice communications,* Report no. DOT/FAA/AM-95/15, Washington, DC: FAA.

[11] Cardosi, K., B. Brett, and S. Han, 1996, *An analysis of TRACON (Terminal Radar Approach Control) controller-pilot voice communications,* Report no. DOT/FAA/AR-96/66, Washington, DC: U.S. Department of Transportation, FAA.

[12] Federal Aviation Administration, 2004, *FAA Order 7110.65P Air Traffic Control, www.faa.gov/ATPUBS/ATC/.*

[13] Baddeley, A., 2000, The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences, 4,* 417-23.

[14] Morrow, D. and Prinzo, O., 1999, *Improving pilot/ATC voice communication in general aviation,* Report no. DOT/FAA/AM-99/21, Washington, DC: FAA.

[15] Krippendorff, K., 1980, *Content analysis: An introduction to its methodology,* Thousand Oaks, CA: Sage Publications.

[16] Rantanen, E. and N. Kokayeff, 2002, Pilot error in copying air traffic control clearances, In *Human Factors and Ergonomics Society 46th Annual Meeting Proceedings,* Santa Monica, CA Civil Aviation Authority, (April 2000).

[17] Cardosi, K., 1999, *Human factors for air traffic control specialists: A user's manual for your brain,* Report no. DOT/FAA/AR-99/39, Washington, DC: FAA.

[18] Cardosi, K., 1993, *An analysis of en route controller-pilot voice communication,*. DOT Report no. DOT/FAA/RD-93/11, Washington, DC: FAA.

[19] Morrow, D. and M. Rodvold, 1993, *The influence of ATC message length and timing on pilot communication,* NASA Contract Report 177621, Moffett Field, CA: NASA ARC.

[20] Morrow, D., M. Rodvold, A. McGann, and M. Mackintosh, 1994, *Collaborative strategies in air-ground communication,* Technical Paper #942138, Los Angeles, CA: Society of Automotive Engineers.

[21] Federal Aviation Administration, 2006, *Aeronautical Information Manual: Official guide to basic flight information and ATC procedures,* www.faa.gov/ATpubs/AIM/.

[22] Air Safety Week, Dec 13, 1999, *Data Link Communications Could Reduce Delays and Improve Safety.*

[23] Prinzo, O. and M. McClellan, 2005, *Terminal radar approach control: Measures of voice communications system performance*, Report no. DOT/FAA/AM-05/19, Washington, DC: FAA.

Keywords:

ATC Communication; Language; Air traffic control; International pilot controller communication.

Biography

O. Veronika Prinzo received a Ph.D. from Kent State University in Experimental Psychology with an emphasis in Psycholinguistics and Cognitive Psychology. She taught for several years in the Great Lakes Consortium of Colleges and is presently an Engineering Research Psychologist at the FAA's Civil Aerospace Medical Institute. Her 14 yrs of aviation research involves the study of pilot-controller communication and data link applications to enhance safety, improve efficiency, and facilitate understanding.

Al and Ruby Hendrix have more than 38 combined yrs of government service. After retiring from the FAA they formed Hendrix Consulting Services. They have assisted in establishing, collecting, assembling, and analyzing research data for 14 yrs in the following fields – Age 60, sleep deprivation/fatigue/effects of shift work, data link, ATC verbal and nonverbal communication, and effectiveness of pilot/controller communication.