Safety Science 73 (2015) 73–79

Contents lists available at ScienceDirect

Safety Science

journal homepage: www.elsevier.com/locate/ssci

Miscommunication in general aviation: The influence of external factors on communication errors

Brett R.C. Molesworth [a,]*, Dominique Estival [b]

[a] School of Aviation, University of New South Wales, NSW, Australia
[b] MARCS Institute, University of Western Sydney, NSW, Australia

ARTICLE INFO

Article history:
Received 11 July 2014
Accepted 7 November 2014

Keywords:
Communication
Errors
Aviation
Safety
Pilot-ATC

ABSTRACT

Miscommunication in aviation remains a serious threat to safety. Factors such as pilots workload, quality of audio signal, accent of pilot or controller, English language proficiency of operator, and failure to use standard phraseology are all thought to contribute to communication errors. Hence, the aim of the present research was to investigate if a relationship existed between four known factors moderating communication and communication accuracy. Seventeen pilots completed a total of eight separate simulated flights (presented in counterbalanced order), which were arranged in four flight pairings and the percentage of accurate transmissions were compared between each flight pairing. The results revealed that requiring four or more items in one radio transmission degraded communication performance. Similar results were noted when pilots were under high workloads. Eliminating prosodic features such as pauses in radio transmissions also increased communication errors; most notably for pilots whose native language was not English. There was no effect of airways congestion on pilot communication performance. The results are discussed from a theoretical and applied perspective.

1. Introduction

"Aviate, navigate, communicate" is an important adage pilots are required to remember. It is therefore unlikely to be fortuitous that miscommunication (i.e., communication errors) features prominently in many aviation accidents. This has not gone unnoticed with renewed emphasis from aviation authorities focusing on aviation terminology as well as language proficiency standards (International Civil Aviation Organization – ICAO, 2007; Moder, 2013), with the ultimate objective of improving safety. However, there are many moderating factors such as pilots' workload, quality of audio signal, accent of pilot or controller, English language proficiency of operator, and failure to use standard phraseology that are likely to contribute to communication errors. Therefore, the main aim of the present research was to investigate if a relationship existed between four known factors moderating communication and communication accuracy.

In general aviation, it is a requirement that all aircraft operating in controlled airspace have a serviceable (i.e., functioning) radio. It is also a requirement that all pilots hold a radio telephony licence. In 2003, the International Civil Aviation Organization introduced minimum levels of English language proficiency for both pilots and air traffic controllers, which came into effect in 2008 within Australia. These requirements, and in particular the latter two, are designed to improve radio transmission skills, and ultimately enhance safety.

The added inclusion of the English language proficiency skills is on top of existing safeguards to protect against communication errors in aviation such as: English as the international language; the use of standard phraseology (e.g., 'roger' and 'wilco' for acknowledgement of instructions); international phonetic alphabet (e.g., Alpha, Bravo, Charlie, Delta, etc.); prescribed pronunciation of letters and numbers (e.g., 'IN dee A' for India, 'wun' for number one, 'nin er' for number nine, one thousand five hundred for 1500); and read-back requirements (e.g., only key elements of the instructions or clearances are required to be read back; Aeronautical Information Publication – AIP, Airservices Australia, 2005).

Notwithstanding these principles, radio transmission skills such as pronunciation, speech rate and accent have been cited as leading contributing factors in communication problems in both commercial aviation and general aviation (EUROCONTROL, 2006; Tiewtrakul and Fletcher, 2010; Estival and Molesworth, 2012). Take call-signs for example, EUROCONTROL (European member state organisation with its central focus on air traffic management) found in a study with 241 airline pilots and air traffic controllers, that twenty per cent of respondents indicated that they experience

---

* Corresponding author at: School of Aviation, Room 205c, UNSW, Sydney, NSW 2052, Australia. Tel.: +61 2 9385 6757; fax: +61 2 9385 6637.
E-mail address: b.molesworth@unsw.edu.au (B.R.C. Molesworth).

http://dx.doi.org/10.1016/j.ssci.2014.11.004
0925-7535/© 2014 Elsevier Ltd. All rights reserved.

communication problem with call signs on a weekly basis. Controller accent was cited as the leading contributing factor (34%) closely followed by controller speech rate (28%), pilot distraction (25%), pilot expectation (22%) and pilot fatigue (20%). Similar findings were evident when respondents were asked about frequency changes: controller accent (51%), controller speech rate (42%) and pilot distraction (43%).

Estival and Molesworth (2009) found similar results when they surveyed 36 pilots from various flight training institutions at Bankstown airport, in Sydney Australia about miscommunication in general aviation. When pilots were asked what they found most challenging in general aviation communication, pilots noted 'understanding other pilots' as most challenging. Subsequent comments from pilots indicated that communicating with non-native English speaking pilots was particularly challenging. In a follow-up study, with 83 pilots from various flight training institutions in both New South Wales (NSW) and the Australian Capital Territory (ACT) in Australia, they found similar results, with 'understanding other pilots as the most challenging aspect of communicating in aviation'. No differences were noted in responses based on native language background i.e., native English speaker (NS) and non-native English speakers (ESL), however understanding non-native English pilots featured predominantly as one issued raised by many pilots. When asked if they had been in a situation where they did not fully understand the instructions from air traffic control, over half of the pilots noted that they had. Estival and Molesworth interpreted these findings as evidence that poor communication skills were a likely factor leading to communication problems.

However, factors generally outside the pilot control can also contribute to miscommunication. For example, Barshi and Farris (2013) found in a pen and paper study with non-pilots that when four or more items were presented in one transmission, communication errors doubled compared to when there were three or fewer items in a transmission. The prosodic features of the message (i.e., intonation, pauses, and stress) may also contribute to communication errors (Estival and Molesworth, 2009). Moreover, presenting information in a transmission without pauses, or any emphasis on important words, is likely to add to communication errors. Pilot workload is also thought to add to communication errors (Lin et al., 2012), with communication itself further adding to workload (Linde and Shively, 1998). Congested radio frequencies are also thought to adversely affect communication performance (Morrow et al., 1993).

The present research attempts to extend the research conducted by EUROCONTROL (2006) and Estival and Molesworth (2009, 2012) with the intent to investigate the impact of factors outside the control of the pilot on communication errors. Specifically, the present research will seek to answer the following four questions:

1. Does the number of items in a transmission, such as four or more items per radio transmission increase pilot communication errors?
2. Do the prosodic features of a message, such as a radio transmission without pauses, increase pilot communication errors?
3. Is there a relationship between pilot workload and pilot communication errors?
4. Does airspace congestion adversely affect pilots' ability to communicate?

2. Method

2.1. Participants

Seventeen pilots (one female), eight of whom were native English speakers (NS) volunteered for the research. The average age of the participants was 30.82 ($SD = 13.97$) years. The native language of the non-native English speakers (English as a second language - ESL) included: Cantonese (4), Chinese (1), Malayalam (1), Italian (1), Danish (1), Russian (1). On average, the ESL speakers reported to have spoken English for 17.11 ($SD = 11.96$; range 2–35) years. The research, including all stimuli was approved in advance by the University of Western Sydney Ethics Committee.

2.2. Design

The study comprised a 2 × 4 mixed repeated measures design with language background as the between-groups factor containing two levels (NS vs. ESL), and flight scenario as the repeated measures factor containing four different flight pairings (Pauses, Information Density, Workload, and Frequency Congestion). The four flight pairings (eight flights in total) were presented in a counterbalanced order as per a 4 × 4 Latin square design. A Latin Square design was chosen, as opposed to a balanced Latin square design, because of the undesirable adjacency which a balanced Latin Square would have given to the two flights in each pair. The dependent variable in all flights was communication accuracy.

The ATC transmissions (calls) played to the pilots were pre-recorded as separate calls according to flight scenarios designed by the two researchers. The scenarios for each flight were recorded separately, with a male aviation professional with more than 30 years of flying experience as the ATCO. The calls for each flight were then concatenated in a single sound file, with time for pilot answers inserted between each ATC call. In total there were 126 transmission opportunities for each pilot throughout the eight different flights.

As can be seen in Table 1, there were four pairs of flights, with one of the test flights acting as the baseline (easy condition) and the other flight in the pair serving as the experimental flight. Three flight scenarios (flight pairings) were navigation flights, with the fourth scenario an approach to land at a local airport.

The first flight pairing compared read-back errors when ATC instructions contained pauses between items (Flight 1A) vs. no pauses between items (Flight 1B), i.e., 'Pause condition'. In other words, in the no pause condition (Flight 1B) the ATC instruction was one continuous utterance (e.g., "ABC Camden Tower Maintain 3500 and maintain 160 Contact Sydney Centre on 124.55").

The second flight pairing compared read-back errors when each ATC transmission contained no more than 3 items (Flight 2A) vs. when ATC instructions contained 4 or more items (Flight 2B), i.e., 'Information Density condition'. For example, compare the ATC transmission "ABC, Sydney Centre. Climb to 4500. Track 250." in Flight 2A, with "ABC, Sydney Centre, climb to 4500. Track 250. Traffic is a Cessna at your 9 o'clock, Report sighted." in Flight 2B.

Table 1
Overview of the four flight pairings.

Flight #	Flight description	Departure point	Destination point	# Of possible pilot transmissions
1A	With pauses	Camden	Wollongong	16
1B	Without pauses	Camden	Wollongong	16
2A	3 or fewer items per transmission	Camden	Goulburn	18
2B	4 or more items per transmission	Camden	Goulburn	20
3A	Low workload	Camden	Canberra	13
3B	High workload	Camden	Canberra	13
4A	No radio congestion	Entry point 2RN	Bankstown	11
4B	High radio congestion	Entry point 2RN	Bankstown	11

The third flight pairing compared read-back errors when participants were under normal workload in a routine navigation flight (Flight 3A) vs. when participants were under high workload after being requested to perform in-flight fuel calculations due to adverse weather conditions in the same navigation scenario (Flight 3B), i.e., 'Workload condition'.

The final pairing compared read-back errors when air traffic in the area was sparse (Flight 4A) vs. when air traffic was congested (Flight 4B), i.e., 'Frequency Congestion condition'. Extra calls were recorded to simulate congested traffic at the airport and inserted between ATC transmissions to pilot flying. The expected pilot transmissions were the same in Flight 4A and Flight 4B.

The dependent variable in all flights was the percentage of correct transmissions to the total number of possible calls per flight. As can be seen in Table 1, this varied as a result of flight pairing.

2.3. Materials

The laboratory equipment comprised: X-Plane 6.21 featuring a Cessna 172 aircraft (with call sign "ABC"), a Personal Aviation Training Devices (PCATD) with one twenty-one inch flat screen monitor, Elite rudder pedals, and two additional computers: one to play the audio stimuli (though an aviation headset) and one to record the pilot's responses (using the Audacity software).

The test documentation comprised: an information sheet, a consent form, and a demographics questionnaire asking participants to provide their age, sex, native language, number of years speaking English and flying experience.

2.4. Procedure

Participants were recruited through two methods. First, flyers advertising the research were placed at a number of different flight training schools at both Bankstown and Camden aerodromes in the Sydney basin, New South Wales (NSW) Australia. Second, students within the Bachelor of Aviation programme at the University of New South Wales were informed about the research during class and invited to place their name and contact details on a participant sign-up sheet. With all pilots, a mutually suitable time was arranged to conduct the research. On the day of the experiment, participants were asked to complete a consent form and a demographics questionnaire (age, sex, language background). They were then given instructions for each flight, consisting of weather forecast and a flight plan. The eight different test flights were presented in a Latin square design.

In order to replicate the applied environment as much as possible, reproduced aircraft noise of a Cessna 172 during cruise at 65 dBA was played throughout each flight. At the conclusion of the session, participants were thanked for their time and given their reimbursement (either $50 cash or a book voucher to equivalent value). The average time to complete the task was 2 h for each pilot. A total of 17 pilots completed the task, for a total number of 136 flights. The data recorded for each flight was of two types, audio and pilot actions (basic flight performance parameters – heading and altitude).

3. Results

3.1. Overview of data analysis

The main aim of the present research was to investigate the relationship between some flight conditions and communication accuracy. Four different flight conditions were of particular interest: speed of ATC transmission, amount of information transmitted, workload and radio frequency congestion. Communications from pilots were recorded and accuracy was analysed for each of the flight pairings, in which one of the flights served as a baseline and the other flight as the experimental condition. Hence, a repeated measures analysis was conducted for each flight pairing. Since it has been previously been shown that it is a factor which affects aviation communication, language background (Jang et al., 2014), that is native English speakers (NS) vs. English as a Second Language speakers (ESL) featured as a between-groups factor IV in all analyses. In addition, and for similar reasons, pilot licensing qualifications, that is low level licence (Private Pilot Licence (PPL) or lower) vs. high level licence (Commercial Pilot Licence (CPL) or higher) served as the second between-groups factor IV in all analyses. Therefore, a series of $2 \times 2 \times 2$ mixed repeated measures ANOVAs were conducted with percentage of accurate transmissions as the sole repeated measures dependent variable in each flight, and language background and licence qualification as the two between-groups factors.

The audio recordings were transcribed and each transmission was coded as 'Correct' or 'Incorrect', given a set of admissible variants for each pilot's read-back or report. For example, one Air Traffic Control (ATC) transmission provided during Flight 1A was "Alpha Bravo Charlie climb to 4500, maintain 160". In response, pilots should reply "Climb to 4500, maintain 160. ABC". Omitting "climb" or "maintain" was accepted, as per the AIP (Airservices Australia, 2005), but any other deviation from this response was coded as incorrect.

Due to technical issues, audio recordings for 6 flights were either lost or incomplete, with eighty-five possible pilot transmissions missing out of a total of 2142 transmission opportunities[1]. As recommend by Tabachnick and Fidell (2013) the entire data set was reviewed independently by two qualified professionals (researchers who both hold a CPL) and the missing data (number of Correct/Incorrect calls) was replaced through a process which involved the comparison of pilots' performance with their performance on other flights and with other pilots' performance across flights (Prior Knowledge Process; Tabachnick and Fidell, 2013).

Pilot radio transmissions occurred during a simulated flight, and specifically the transmissions from ATC instructed pilots to change heading, altitude, and radio frequency. Therefore it was important to ensure that pilots did not neglect their flying duties while communicating and the pilots' flight performance was monitored so as to exclude pilots who failed to follow ATC instructions. No pilots failed to follow ATC instructions and hence none were excluded.

3.2. Pauses between items in ATC transmissions (Flights 1A and 1B)

Throughout each of the two flights (with pauses vs. no pauses between items in ATC transmissions), there were a total of 16 expected pilot transmissions (read-backs or reports). In order to determine if performance in terms of percentage of correct transmissions varied between the two different flight conditions, a mixed repeated measures analysis of variance (ANOVA), with native language as one of the between-groups factors and licence type as the second between-groups factor was employed. With the ANOVA test assumptions satisfactory, and alpha set at .05, the results failed to reveal a main effect for flight condition, or an interaction between flight condition and language background and between flight condition and licence type, largest F, $F(1,13) = .537$, $p = .477$, $\eta_p^2 = .040$. There was however a three-way interaction between flight condition, language background and licence type, $F(1,13) = 5.797$, $p = .032$, $\eta_p^2 = .308$.

---

[1] Flight 1A, Participant 12, 6 transmissions; Flight 1B, Participant 15, 16 transmissions; Flight 2A, Participant 12, 18 transmissions; Flight 3A, Participant 12, 17 transmissions and Participant 13, 17 transmissions; Flight 4A, Participant 12, 11 transmissions.

The three-way interaction was analysed by simple effect analyses performed separately for Flights 1A and 1B using percentages of correct transmissions, with a Bonferroni adjusted alpha of .025 because of the two simple effects. In view of the small sample sizes, the non-parametric equivalent to the one-way ANOVA, namely Mann–Whitney *U* test, was employed. As can be seen in Fig. 1, notable performance differences were evident during Flight 1B for the ESL pilots only. Specifically, the Mann–Whitney nonparametric test indicated a marginally significant difference between groups based on flying qualifications (low pilot qualification vs. high pilot qualification), $z\ (N = 9) = 1.99$, $p = .046$. This result suggests that ESL pilots with low levels of pilot qualification found it more difficult to communicate accurately during the flight with no pauses in ATC communications than the other pilots.

The main effect for language background was significant, $F(1,13) = 15.56$, $p = .002$, $\eta_p^2 = .545$, and there was a main effect for pilot qualification, $F(1,13) = 4.746$, $p = .048$, $\eta_p^2 = .267$, as well as an interaction between language background and pilot qualification, $F(1,13) = 4.746$, $p = .048$, $\eta_p^2 = .267$. The percentage of correct transmissions for the NS group was 65.63 ($SE = 3.08$) while the percentage of correct transmissions for the ESL group was 49.30 ($SE = 2.53$). This result suggests that ESL pilots found communicating more problematic in that condition, and as a result committed more communication errors than NS pilots.

The same was true for licence type. Pilots with a CPL or higher qualification had a greater number of accurate transmissions (62.11, $SE = 2.43$) compared to pilots with a PPL or lower pilot qualification (53.44, $SE = 3.15$). This suggests that pilots holding a PPL or lower qualification found it more difficult to communicate accurately in that condition compared to pilots with a CPL or higher qualification (see Tables 2 and 3).

Since there was no main effect between flight condition and pilot qualification, two separate correlational analyses were performed, one for Flight 1A and one for Flight 1B, comparing total flying experience and communication accuracy. The experience of pilots in the study varied from ab initio pilots with as little as 42 flight hours to experienced pilots with as much as 3500 flight hours. The results of the two Pearson product–moment correlation failed to reveal a relationship between flight experience and communication accuracy, largest $r(17) = -.346$, $p = .173$. This result suggests that what is important in terms of communication is the level of training received, as opposed to the number of hours flown.

3.3. Number of items per ATC transmission (Flights 2A and 2B)

Since the aim of this flight pairing was to examine the impact of the number of items per transmission (three or fewer items per transmission vs. four or more items per transmission), hence complexity of ATC transmissions on communication accuracy, the second flight required two more transmissions from the pilots, therefore, the total number of transmissions for Flight 2A was 18, and 20 for Flight 2B. Consistent with the previous analysis, percentages of correct transmissions from the two different flights were compared using a mixed repeated measures analysis of variance (ANOVA), with native language and licence qualification as the two between-groups factors. With the ANOVA test assumptions satisfactory, and alpha set at .05, the results revealed a main effect for flight condition, $F(1,13) = 7.21$, $p = .019$, $\eta_p^2 = .357$. No other main effects or interactions were significant, largest *F*, $F(1,13) = 4.076$, $p = .065$, $\eta_p^2 = .239$. The percentage of correct transmissions when the ATC calls contained three or fewer items (F2A) was 58.91 ($SE = 4.46$), however when there were four or more items per ATC call (F2B), the percentage of correct pilot transmission fell to 40.75 ($SE = 5.24$; see Tables 4 and 5). These results suggest that, irrespective of native language background or flight qualification, increasing the number of items required in one radio communication significantly affected pilots' ability to accurately respond, thus confirming results by Barshi and Farris (2013).

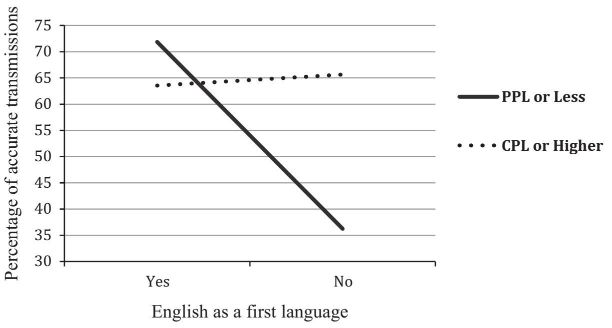

Fig. 1. Percentage of accurate transmissions for Flight 1B, distributed across native language background (NS vs. ESL) and licence type (PPL or lower vs. CPL or higher).

Table 2
Percentage of correct transmissions for NS and ESL speakers distributed across the two flights in flight pairing 1 (pause vs. no pauses in ATC transmissions).

Flight condition	Native language background			
	ESL ($n = 9$)	*SD*	NS ($n = 8$)	*SD*
F1A (pauses)	48.61	8.72	65.63	10.56
F1B (no pauses)	49.31	21.30	65.63	6.68

Table 3
Percentage of correct transmissions for licence type (PPL or lower vs. CPL or higher) distributed across the two flights in flight pairing 1 (pause vs. no pauses in ATC transmissions).

Flight condition	Licence type			
	PPL or lower ($n = 7$)	*SD*	CPL or higher ($n = 10$)	*SD*
F1A (pauses)	50.00	10.21	61.25	12.77
F1B (no pauses)	46.43	20.68	64.38	11.43

The results of two separate Pearson product–moment correlations also failed to reveal a relationship between flight experience (as measured by the number of flying hours) and communication accuracy on either Flight 2A or 2B, largest $r(17) = .324$, $p = .204$. This result reflects the correlational analysis from the first flight pairing (1A and 1B) and suggests that increased flying experience did not significantly improve communication accuracy and did not help pilots when the number of items per transmission was too great.

3.4. Workload (Flights 3A and 3B)

Consistent with the two previous analyses, percentages of correct transmissions from the two flights, i.e., under low workload (F3A) and high workload (F3B), were compared using a mixed repeated measures analysis of variance (ANOVA), with native lan-

Table 4
Percentage of correct transmissions for NS and ESL speakers distributed across the two flights in flight pairing 2 (3 or fewer items vs. 4 or more items).

Flight condition	Native language background			
	ESL ($n = 9$)	*SD*	NS ($n = 8$)	*SD*
F2A (3 or fewer item)	50.62	22.29	65.28	13.53
F2B (4 or more items)	36.67	24.75	46.88	13.61

Table 5
Percentage of correct transmissions for licence type (PPL or lower vs. CPL or higher) distributed across the two flights in flight pairing 2 (3 or fewer items vs. 4 or more items).

Flight condition	Licence type			
	PPL or lower ($n = 7$)	*SD*	CPL or higher ($n = 10$)	*SD*
F2A (3 or fewer item)	46.83	21.24	65.00	15.28
F2B (4 or more items)	30.71	13.36	49.00	21.58

Table 7
Percentage of correct transmissions for licence type (PPL or lower vs. CPL or higher) distributed across the two flights in flight pairing 3 (low workload vs. high workload).

Flight condition	Licence type			
	PPL or lower ($n = 7$)	*SD*	CPL or higher ($n = 10$)	*SD*
F3A (low workload)	56.04	14.54	59.23	17.78
F3B (high workload)	40.66	19.72	53.08	8.46

guage and licence qualification as the two between-groups factors. Since the request for fuel calculation occurred after the fourth ATC call, data was analysed following this point (giving a total of 13 possible transmissions out of 17 for the whole flight). With the ANOVA test assumptions satisfactory, and alpha set at .05, the results revealed a main effect for flight condition, F, $F(1,13) = 6.90$, $p = .021$, $\eta_p^2 = .368$. There was however no interaction between flight condition and language background, flight condition and licence qualification, or flight condition, language background and licence qualification, largest F, $F(1,13) = .79$, $p = .39$, $\eta_p^2 = .057$. The percentage of correct transmissions during Flight 3A (low workload) was 60.22 ($SE = 4.21$) while the percentage of correct transmissions during Flight 3B (high workload) was 49.58 ($SE = 2.92$). This result suggests that increasing the pilots' workload adversely affected their ability to communicate effectively during flight.

There was a main effect for language background, F, $F(1,13) = 6.39$, $p = .025$, $\eta_p^2 = .2301$; and no other interactions were significant, largest F, $F(1,13) = 3.56$, $p = .082$, $\eta_p^2 = .215$. The percentage of correct transmissions for NS pilots was 62.50 ($SE = 4.64$) while the percentage of correct transmissions for ESL pilots was 47.31 ($SE = 3.82$; see Tables 6 and 7). Consistent with the results from flight pairing 1A and 1B, this result shows that pilots with native English language background found it easier to communicate than to ESL pilots, irrespective of workload.

As a manipulation check to ensure that performance between the two flights was not affected by other unknown or uncontrolled factors, the number of correct transmissions prior to the increase in workload (after 4 ATC calls) was compared between the two flights. With assumptions of normality met, the results of a dependent t test failed to reveal a statistically significant difference between correct transmissions in Flight 3A and in Flight 3B, $t(17) = 1.23$, $p = .236$. This result suggests that pilots found the two flights similar up to the point when they were asked to increase their workload in Flight 3B.

The results of two separate Pearson product–moment correlations also failed to reveal a relationship between flight experience and communication accuracy in either of the two flights, largest $r(17) = .138$, $p = .598$. Again this result suggests that a higher number of flying hours did not lead to an improvement in communication accuracy.

3.5. Congested radio frequency (Flights 4A and 4B)

In this flight pairing, the total number of transmissions expected from the pilot was 11 for each flight, but there were additional ATC calls directed at other aircraft and a number of transmissions from other traffic on the same radio frequency in Flight 4B. The series of analyses for Flights 4A and 4B followed a similar format as the other three flight pairings, namely percentages of correct transmissions from the two different flights, i.e., uncongested frequency (F4A) and congested frequency (F4B), were compared between the two flights using a mixed repeated measures analysis of variance (ANOVA), with native language and flight qualification as the two between-groups factors. With the ANOVA test assumptions satisfactory, and alpha set at .05, the results failed to reveal any main effect or interaction, largest F, $F(1,13) = 3.05$, $p = .104$, $\eta_p^2 = .190$ (see Tables 8 and 9). This result suggests that radio frequency congestion does not affect pilots' ability to communicate effectively, irrespective of their language background or pilot qualification.

The final two correlational analyses also failed to reveal a relationship between flight experience and communication accuracy on either of the two different flights, largest $r(17) = .250$, $p = .333$. Again this result suggests that more flying experience did not lead to improved communication accuracy irrespectively of radio frequency congestion. In other words, pilots of all levels of qualification and experience are able to ignore chatter on the radio and to make their calls appropriately.

3.6. Synopsis across all four flights

The experimental design allowed for direct comparisons within each flight pairing (e.g., between F1A and F1B), but did not permit direct comparison between the four different scenarios (e.g., F1A and F1B compared to F2A and F2B). Although the first three flight scenarios all commenced from Camden Aerodrome and tracked in a southerly direction to Wollongong, Goulburn or Canberra, this is where the similarities cease. The different navigation scenarios contained different air traffic control instructions for obvious reasons, namely to reduce as much as possible the potential of a learning effect, and flight scenario number four required participants to approach and land at Bankstown Aerodrome. Nevertheless, the experimental design still permits some comparisons between the four different flight scenarios, which provide some interesting observations. For example, data presented in Tables 8 and 9 illustrates that, irrespective of language background, pilots seemed to find communicating during flight scenario number 4 (approach and landing at Bankstown Aerodrome) less challenging than the other three flight scenarios. On average, 68% of their communications were correct in Flights 4A and 4B. In contrast, just over half (53%) of pilot communications during the other six flights (tracking south to Wollongong, Goulburn or Canberra) were correct. Since scenario four, i.e., approaching an aerodrome to land (even while

Table 6
Percentage of correct transmissions for NS and ESL speakers distributed across the two flights in flight pairing 3 (low workload vs. high workload).

Flight condition	Native language background			
	ESL ($n = 9$)	*SD*	NS ($n = 8$)	*SD*
F3A (low workload)	53.85	14.90	62.50	17.17
F3B (high workload)	39.32	15.12	57.69	7.12

Table 8
Percentage of correct transmissions for NS and ESL speakers distributed across the two flights in flight pairing 4 (no radio congestion vs. high radio congestion).

Flight condition	Native language background			
	ESL ($n = 9$)	*SD*	NS ($n = 8$)	*SD*
F4A (no radio congestion)	61.62	18.62	76.14	13.69
F4B (high radio congestion)	66.67	14.38	70.45	16.65

Table 9
Percentage of correct transmissions for licence type (PPL or lower vs. CPL or higher) distributed across the two flights in flight pairing 4 (no radio congestion vs. high radio congestion).

Flight condition	Licence type			
	PPL or lower ($n = 7$)	*SD*	CPL or higher ($n = 10$)	*SD*
F4A (no radio congestion)	59.74	14.71	74.55	17.56
F4B (high radio congestion)	72.73	10.50	65.45	17.57

ignoring chatter over the radio), could be considered the most routine of all the four scenarios, this result is hardly surprising, nonetheless noteworthy.

Comparing the three southerly cross country flights under adverse conditions, i.e., Flight 1B with no pauses between items in ATC transmissions, Flight 2B with more than four items per ATC transmission, and Flight 3B under high workload, the flight condition that yielded the worst performance amongst both NS and ESL pilots was Flight 2B (four items per ATC transmission). This result is consistent with Barshi and Farris (2013) and suggests that having a large number of items in ATC transmissions (more than four) is particularly problematic for effective communication.

In terms of language background and performance, the results largely failed to reveal any surprises, with NS pilots producing more correct transmissions as a whole (63%) compared to ESL pilots (51%). What is surprising however, is the large number of errors in radio transmissions by pilots in the first place. For NS pilots, the percentage of incorrect transmissions was approximately 40%, while for ESL pilots this approached 50%. Therefore, the combined accuracy of content communicated over the radio in GA is a little more than fifty per cent. This result would suggest either that aviation communication is very difficult or that GA pilots are poorly trained when it comes to radio communication. It seems likely that aviation communication is in fact harder than generally considered, and this in itself would indicate that the quality of the training needs to be improved.

4. Discussion

The main aim of the present research was to investigate the relationship between flight condition and communication accuracy by pilots across four different flight scenarios. In addition, the present research sought to investigate whether known factors such as native language background, pilot qualification or flight experience interacted with pilots' ability to communicate accurately during flight. The results revealed a complex relationship between these variables. For example, and not surprisingly, overall NS pilots committed fewer errors than ESL pilots in their radio transmissions. However, when the task was very familiar such as arriving at the local aerodrome (i.e., Flight 4B) or when the radio transmissions were particularly difficult (i.e., Flight 2B), both NS and ESL pilots found it equally easy or hard. In contrast, when the task was less familiar, ESL pilots committed more communication errors than NS pilots. This result would seem to suggest that the act of communicating over the radio is cognitively taxing, and when combined with other high workload tasks, performance deteriorates. For pilots, this may come as no surprise as it highlights the importance of the aforementioned adage: aviate, navigate, communicate.

What is of particular concern is the overall poor communication performance of all pilots. Given that it is widely reported that miscommunication is a contributing factor in many aviation incidents and accidents (Cushing, 1994; Corradini and Cacciari, 2002), the results from the present research suggest that it is an area that warrants particular attention. The results further suggest that relying on the natural development of communication skills as a progression of flight experience is not an option; recall licence qualification and not hours of flight experience yielded differences in communication performance. In the context of aviation, and given that theory tests for flying qualifications such as PPL, CPL and Air Transport Pilot Licence (ATPL) require pilots to achieve a minimum of 70% (80% for ATPL or CPL Air Law), correctly communicating in approximately only half of all the transmissions, as in the present study, is alarming.

In relation to the four research questions at the centre of this study, the results suggest that the absence of pauses in transmission is particularly problematic for ESL pilots with low pilot qualifications (question 1). Question two was concerned with the number of items presented in each radio transmission, and the results are definitive. Including four or more items in a radio transmission adversely affects communication accuracy and native language or pilot qualification cannot compensate for the increased level of difficulty. Pilot workload was also found to adversely affect communication performance (question 3), however this time having a native language background was sufficient to offset this effect. Finally, the results from the present research suggest that all pilots, including low qualification pilots and ESL pilots were able to filter out any radio transmissions that were not relevant to them (question 4). This raises the question whether pilots were actually filtering, or ignoring, them; if the latter, this has important safety implications and warrants further investigation.

4.1. Limitations and future research

While the results from the present research clearly highlight the effects of various conditions of pilots' communication performance, they need to be interpreted within the confines of the study. The present study drew upon pilots working or being trained within the general aviation sector of aviation. These pilots were either predominately based or regularly flew to or from one of the busiest general aviation airports in the southern hemisphere (Sydney Metro Airports, 2014), which functions as a controlled aerodrome extending well into the night. How pilots from other sectors of aviation such as commercial aviation or military aviation, or pilots not normally exposed to controlled airspace perform remains unknown; hence an area for future research.

Given that the results from flight pairing three (workload) suggests that communicating in aviation is cognitively taxing (in line with Linde and Shively, 1998) it would be prudent to examine this in greater detail in order to determine which aspect of this task are taxing (i.e., remembering what has to be said, order in which items need to be placed, or noise; Molesworth and Burgess, 2013), and to investigate training methods to reduce their effect. In addition, and for the purpose of highlighting how cognitively taxing communication in aviation is, future research should attempt to quantify this. In doing so, this would highlight the challenges pilots face, and in particular ESL pilots. This in turn should assist in deriving countermeasures to mitigate their effects.

Future research should also examine the relationship between communication and flight performance. In the present study a flight simulator was employed to reflect the operational environment and an exclusion criteria was set where any pilot who failed to follow ATC instructions was excluded (no pilots excluded). What remains an area for future research is to collect and scrutinise the pilots' flight performance, such as their ability to hold and maintain altitude and heading, in order to determine if there is a trade-off between flight performance and communication (i.e., "Aviate > Navigate > Communicate"). Finally, having established the effect of external factors on communication errors, future research should examine which aspect of the radio transmission pilots found most challenging.

5. Conclusion

Aviation is a high hazard industry. However, it remains one of the safest modes of transportation, largely due to the ingenuity of design and the resilience of key personnel such as pilots. That said, improvements in safety can always be achieved and communication is one area requiring particular attention. The present research highlights that for ESL pilots, communication remains a challenge in aviation. However, if the flying operational requirements become too demanding, even high English language proficiency is not enough to guarantee accurate communication. Hence it is clear from the present research that greater focus needs to be directed towards improving pilot communication skills.

Acknowledgements

The authors would like to thank Mr. David Maddock for the time spent recording the ATC messages, Mr. Raymond Jang for assisting with the collection of data, and Ms. Annie Zhou for the transcriptions of communication recordings. The authors would also like to thank those pilots who volunteered their time to complete the research.

References

Airservices Australia, 2005. Aeronautical Information Publication. Canberra, Australia (Author).

Barshi, I., Farris, C., 2013. Misunderstanding in ATC Communication: Language, Cognition, and Experimental Methodology. Ashgate, Surrey, England.

Corradini, P., Cacciari, C., 2002. The effect of workload and workshift on air traffic control: a taxonomy of communication problems. Cogn. Technol. Work 4 (4), 229–239.

Cushing, S., 1994. Fatal Words: Communication Clashes and Aircraft Crashes. University of Chicago Press, The Chicago, IL.

Estival, D., Molesworth, B.R.C., 2009. A study of EL2 pilots radio communication in the general aviation environment. Aust. Rev. Appl. Linguist. 32, 24.1–24.16.

Estival, D., Molesworth, B.R.C., 2012. Radio miscommunication: EL2 pilots in the Australian general aviation environment. Linguist. Hum. Sci. 5 (3), 351–378.

EUROCONTROL, 2006. Air Ground Communication Safety Study: Causes and Recommendations (Report # DAP/SAF 2006-09). Brussels, Belgium (Author).

ICAO, 2007. Resolutions adopted by the Assembly – 36th Session. Montreal, Canada (Author).

Jang, R., Molesworth, B.R.C., Burgess, M., Estival, D., 2014. Improving communication in general aviation through the use of noise cancelling headphones. Saf. Sci. 62, 490–504.

Lin, C.J., Lin, P.-H., Chen, H.-J., Hsieh, M-C., Yu, H-C., Wang, E.M-Y., Ho, H.L.C., 2012. Effects of controller-pilot communication medium, flight phase and the role in the cockpit on pilots' workload and situation awareness. Saf. Sci. 50, 1722–1731.

Linde, C., Shively, R.J., 1998. Field study of communication and workload in police helicopter: implications for A1 cockpit design. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting. Human Factors and Ergonomics Society, pp. 237–241.

Moder, C.L., 2013. Aviation English. In: Paltridge, B., Starfield, S. (Eds.), The Handbook of English for Special Purposes. Wiley-Blackwell, Boston, MA, pp. 227–242.

Molesworth, B.R.C., Burgess, M., 2013. Improving intelligibility at a safety critical point: in flight cabin safety. Saf. Sci. 51, 11–16.

Morrow, D., Lee, A., Rodvold, M., 1993. Analysis of problems in routine controller-pilot communications. Int. J. Aviat. Psychol. 3 (4), 285–302.

Sydney Metro Airports, 2014. Sydney Metro Airport Ranks 26th Busiest Airport. <http://www.bankstownairport.com.au/Press_Office/Press_Releases/Press_Releases_2009/Sydney_Metro_Airport_Ranks_26th_Busiest_Airport.aspx> (viewed 09.07.14).

Tabachnick, B.G., Fidell, L.S., 2013. Using Multivariate Statistics. Pearson Education, Boston, MA.

Tiewtrakul, T., Fletcher, S.R., 2010. The challenges of regional accents for aviation English language proficiency standards: a study of difficulties in understanding air traffic control-pilot communications. Ergonomics 53 (2), 229–239.