


## Specificity, Authenticity, and Inseparability: Assessing Integrated Oral Skills for Pilots and Air Traffic Controllers

**Dan DOUGLAS**

Iowa State University

E-mail: [dandoug@iastate.edu](mailto:dandoug@iastate.edu), 

**Abstract:** Because the stakes are so high when assessing the oral language skills of pilots and air traffic controllers, they deserve to be tested in the use of the most specific language features of aviation English and in the most authentic context of their professions. In this talk I will discuss the problem of defining language specificity for professional purposes, the problem of reproducing authentic contexts in assessments, and the problem of the inseparability of language knowledge and professional background knowledge. In terms of specificity, I will talk about “Codes of Relevance,” which relate language use to specific professional contexts. The notion of authenticity is paramount since, unless pilot and ATC test takers can relate to the professional context in a realistic manner, then the judgments we make concerning their language abilities may not be valid. Finally, I will discuss the consequences of the notion that language knowledge and professional background knowledge are theoretically and practically inseparable.

**Keywords:** specificity, authenticity, inseparability, construct, context

### Introduction

When we assess the oral skills of pilots and air traffic control officers (ATCOs) we are challenged to collect performance data of the test takers’ speaking and listening in real time, using English for very specific aviation purposes, in contexts related to ground and in-flight communication, employing simultaneous knowledge of both aviation English and aviation operations. These aspects of the assessment of aviation English – specificity, authenticity, and inseparability – are thus integrated in a complex performance from which we language assessors must derive a measurement of aviation English ability. In this summary I will discuss the notions of specificity – how specific is specific in ESP? – authenticity – how appropriate is the test takers’ response to the communicative situation in the test? – and inseparability – how interconnected are knowledge of the specific purpose language and operational knowledge of aviation practice?

### 1. Specificity

The first question we need to ask regarding the notion of specific purpose language is how specific is specific? Specificity is clearly not an all or nothing proposition: there are differing degrees of specificity, depending on who is doing the communicating, what they’re communicating about, where the communication is taking place, and other contextual factors in the language use situation. In a specific purpose language test, such as tests of aviation Eng-

lish for pilots or ATCOs, we need to take the level of specificity into account by incorporating contextual features into the test format. More than two decades ago, I defined a specific purpose language test as follows:

A specific purpose language test is one in which the test content and methods are derived from an analysis of the characteristics of the specific target language use situation, so that test tasks and content are authentically representative of the target situation, allowing for an interaction between the test-taker's language ability and specific purpose content knowledge, on the one hand, and the test task, on the other (D. Douglas 2000: 19).

Notice that this definition refers to all three of the issues that this paper discusses: specificity, authenticity, and the nature of the relationship between language ability and specific purpose content knowledge (inseparability). Specificity involves the real-world domain of language use, or context; authenticity concerns the distinction between the real-world language use domain and the specific purpose test domain; and the relationship between the components of ability or knowledge measured by the test. In fact, C. Elder (2001) noted the three principal problematic areas identified in my work and recasts them as boundary issues:

- The boundary between and within real world domains of language (Specificity)
- The boundary between the test and the non test situation (Authenticity), and
- The boundary between the components of [language] ability or [content] knowledge measured by the test (Inseparability)

She argues that it is very difficult to determine with any precision where the boundaries lie between these components of specific purpose language ability. However, in a recent publication, the first major treatment of specific purpose language assessment since D. Douglas 2000, U. Knoch and S. Macqueen (2020) propose the concept of *Codes of Relevance* to help us clarify the boundaries somewhat more precisely. Knoch and Macqueen take the *workplace* as the overarching context of specific purpose language use and help us define the boundaries between the language *codes* relevant to subcontexts of language use within the workplace. They exemplify this concept in the following figure:

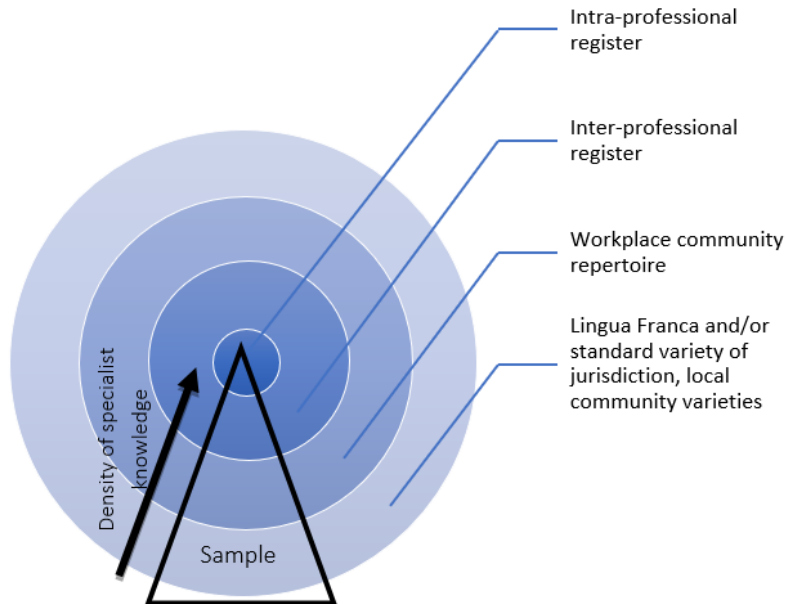


Figure 1. Potential sample range from LAPP [Language Assessment for Professional Purposes] Codes of Relevance (U. Knoch / S. Macqueen 2020: 61).

The concentric circles represent the (admittedly somewhat fluid) boundaries between the domains of language specificity. As one moves in from the outermost circle, the more specific and technical the language becomes and the more specific purpose technical knowledge is required to participate in the communication, until we reach the innermost circle which, in the case of aviation English, might represent a pilot and ATCO in routine highly technical communication. Following are some examples of aviation English communication at each of the four levels of Codes of Relevance.

*a. Intra-professional register.*

Pilot: Seattle Approach, Cherokee 8121K, 20 southwest of Seattle VOR at 7,500, en route Arlington, request transit Class Bravo airspace.

ATCO: Cherokee 8121K, Seattle Approach, squawk 3121 and ident.

Pilot: Squawk 3121, Cherokee 8121K.

ATCO: Cherokee 21K, radar contact 18 miles southwest of SeaTac, Seattle altimeter 29.88, cleared through Class Bravo direct Arlington, descend and maintain 5,500, maintain VFR.

Pilot: Cleared through Class Bravo direct Arlington, descend and maintain 5,500, maintain VFR, Cherokee 21K.

<<http://www.tssflyingclub.org/documents/SampleRadioCalls.pdf>>

In this example, the pilot and ATCO are engaged in highly technical standardized phraseology. The pilot states the current position of the aircraft and requests permission to enter airspace near Seattle-Tacoma International Airport in transit to Arlington Airport. The ATCO acknowledges the position of the aircraft, notes that there is radar contact, grants

permission for the airspace transit, and orders the pilot to descend to an altitude of 5,500 feet and maintain visual flight rules. In the last statement, the pilot acknowledges all this and signs off. A great deal of language and professional content knowledge is required to comprehend and participate in this exchange – a person not familiar with aviation would be able to make little sense of it.

*b. Inter-professional Register*

Pilot: Mayday Mayday Mayday. uh this is uh Cactus fifteen thirty nine - hit birds, we've lost thrust in both engines we're turning back towards LaGuardia.

ATCO: ok uh, you need to return to LaGuardia? turn left heading of uh two two zero.

Captain: two two zero.

First Officer: if fuel remaining, engine mode selector, ignition. ignition.

Captain: ignition.

FO: thrust levers confirm idle.

Captain: idle.

FO: airspeed optimum relight. three hundred knots. we don't have that.

Captain: we don't.

ATCO: Cactus fifteen twenty nine, if we can get it for you do you want to try to land runway one three?

FO: if three nineteen-

Captain: we're unable. we may end up in the Hudson.

<<https://tailstrike.com/database/15-january-2009-us-airways-1549/>>

This is an excerpt from the well-known episode of the ditching of U.S. Airways Flight 1539 into the Hudson River in New York City in January 2009. The captain is Chesley “Sully” Sullenberger, and the First Officer is Jeffery B. Skiles. Shortly after take-off the plane struck a flock of birds and both engines lost power. The captain, first officer, and ATCO are discussing the situation and possible options using “plain language” rather than standardized phraseology as this is an emergency and standardized phraseology will not suffice. The language is still quite technical, though less so than in the previous example. A listener with no knowledge of aviation would be able to make sense of several aspects of the communication.

*c. Workplace Community Repertoire*

Pilot to passengers: “Folks, we have reached our cruising altitude now, so I am going to switch the seat belt sign off. Feel free to move about as you wish, but please stay inside the plane till we land ... it's a bit cold outside, and if you walk on the wings, it affects the flight pattern”.

<[http://www.b737.org.uk/flight\\_attendant\\_humour.htm](http://www.b737.org.uk/flight_attendant_humour.htm)>

In this example of a pilot's humorous announcement to passengers, we see that there is still some technical language being used, e.g., *cruising altitude*, *flight pattern*, but though the announcement is an official workplace communication, the language is mainly colloquial and informal, understandable by most listeners.

*d. Lingua Franca and/or standard variety*

Pilot: So what are you doing during the holidays? Any plans?

FO: Yes. Actually, I'm going on vacation. I'm going to Thailand for two weeks with my parents.

Pilot: That's great! Have you been to Thailand before?

FO: I've flown in there a few times, but it'll be my first extended stay and the first time for my parents.

Pilot: Whoa. How do they feel? I remember taking my parents with me to Mexico. It was a little stressful. Parents aren't easy to travel with sometimes.

FO: My dad's a nervous traveler but my mom is easy going. As long as they have Google Translate and Google Maps, they'll be okay.

<Adapted from <https://helenadailyenglish.com/real-english-conversation-holidays.htm>>

We can imagine this conversation taking place in the cockpit while the pilot and first officer are waiting to begin the pre-flight checks. They are using the standard variety of English and no technical language or background knowledge is required, except perhaps with regard to *Google Translate and Maps*.

*Value of the Codes of Reference Model*

I believe this model can be very helpful to us in the development of tests of aviation English. It can help us match the level of language specificity to the test takers' communicative needs. As we have seen in the examples above, pilots and ATCOs do require a range of language specificity in their workplaces, from situations requiring a high level of specific purpose language ability and associated technical knowledge to casual conversation requiring little of either. The model encourages us to consider the assessment purpose and community of users more fully. Are we assessing aviation English for the purpose of operational flight safety or for ab initio pilots entering training? Who are the ATCOs being assessed communicating with: pilots in aircraft on the ground preparing for takeoff, pilots cleared for departure, pilots enroute, or pilots arriving and taxiing to an arrival gate, or indeed to other ATCOs?

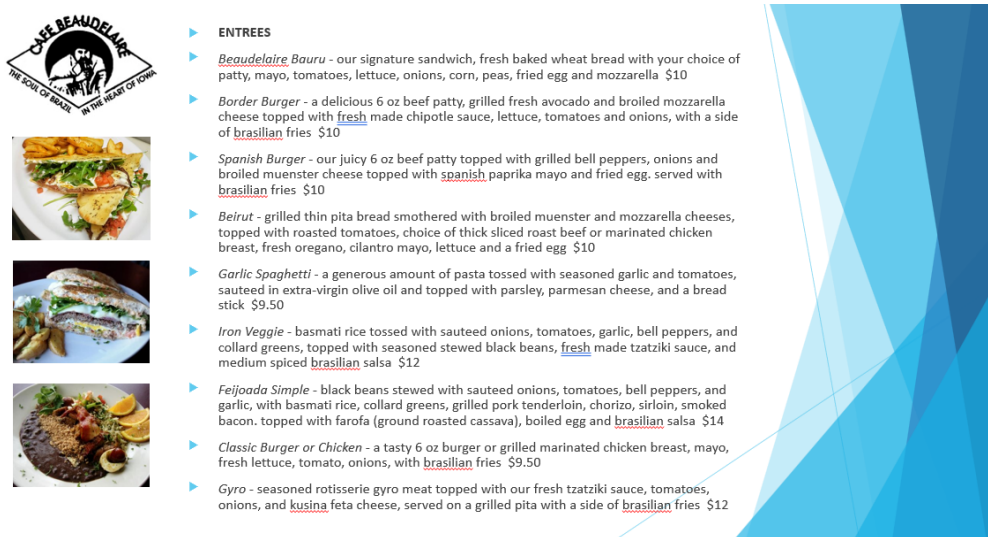
The use of a single variety in a language test underrepresents the dynamism and intercultural character of most aviation language use situations, and the use of the Codes of Relevance model helps us keep tabs on what level of language and technical knowledge we are working with in different parts of the test. Finally, the model allows us to see the relationship between professional registers (or 'professional insider' systems of meaning) and societal dialects: aviation English does not consist of a single register but rather a range of language varieties focused on safe flight operations.

## **2. Authenticity**

Early discussions of authenticity were focused mainly on input materials and the question of whether these should, in the interests of naturalness, be used in unaltered form on a test (C. Elder 2015). It was apparently assumed that merely placing texts generated in the target language use (TLU) situation in the test would ensure that test takers would be interacting with the texts in a manner similar to that in the TLU. It soon became clear, however, that this was not the case. H. Widdowson (1979) was perhaps the first to point out that it was not

the nature of the texts that was at issue but rather whether they were put to use in a manner consonant with the author's intentions. So, if the text involved an instruction from an ATCO to a pilot, e.g., "descend and maintain 5,500, maintain VFR," the issue was how the test taker responded to the input rather than the nature of the input itself. H. Widdowson (ibid) proposed a useful distinction between genuine and authentic input texts. The notion of genuineness is a property of a spoken or written text and results from the text having been produced in an actual communicative situation. Thus, a restaurant menu would be a genuine text in the restaurant, and the diners would no doubt view it as such – unless they suspected they were being duped by being given a fake one! They would also no doubt consider it authentic, since they would respond to the menu in an appropriate manner, using it to order a pleasant meal.

Authenticity thus being a quality associated with appropriate response to a piece of communication (H. Widdowson, 1979). What is key to test takers (or hungry diners) responding to a genuine text in an authentic manner is the features of context associated with the text and the expected response. Contextual features contributing to authenticity include: the setting (e.g. the restaurant), participants (server, diners), purpose (ordering a meal), act sequence (first order drinks, then an appetizer, a main course, and dessert), tone (polite, transactional, welcoming on the part of the server, direct), language (face-to-face, standard English), norms (customer – serving staff local social norms), and genre ("menu") (D. Hymes, 1974).



**CAFE BEAUDELAIRE**  
THE SOUL OF BEAT  
BY THE HEART OF JAZZ

**ENTREES**

- ▶ *Beaudelaire Bauru* - our signature sandwich, fresh baked wheat bread with your choice of patty, mayo, tomatoes, lettuce, onions, corn, peas, fried egg and mozzarella \$10
- ▶ *Border Burger* - a delicious 6 oz beef patty, grilled fresh avocado and broiled mozzarella cheese topped with fresh made chipotle sauce, lettuce, tomatoes and onions, with a side of brasillian fries \$10
- ▶ *Spanish Burger* - our juicy 6 oz beef patty topped with grilled bell peppers, onions and broiled muenster cheese topped with spanish paprika mayo and fried egg. served with brasillian fries \$10
- ▶ *Beirut* - grilled thin pita bread smothered with broiled muenster and mozzarella cheeses, topped with roasted tomatoes, choice of thick sliced roast beef or marinated chicken breast, fresh oregano, cilantro mayo, lettuce and a fried egg \$10
- ▶ *Garlic Spaghetti* - a generous amount of pasta tossed with seasoned garlic and tomatoes, sauteed in extra-virgin olive oil and topped with parsley, parmesan cheese, and a bread stick \$9.50
- ▶ *Iron Veggie* - basmati rice tossed with sauteed onions, tomatoes, garlic, bell peppers, and collard greens, topped with seasoned stewed black beans, fresh made tztziki sauce, and medium spiced brasillian salsa \$12
- ▶ *Feijoada Simple* - black beans stewed with sauteed onions, tomatoes, bell peppers, and garlic, with basmati rice, collard greens, grilled pork tenderloin, chorizo, sirloin, smoked bacon. topped with farofa (ground roasted cassava), boiled egg and brasillian salsa \$14
- ▶ *Classic Burger or Chicken* - a tasty 6 oz burger or grilled marinated chicken breast, mayo, fresh lettuce, tomato, onions, with brasillian fries \$9.50
- ▶ *Gyro* - seasoned rotisserie gyro meat topped with our fresh tztziki sauce, tomatoes, onions, and kusina feta cheese, served on a grilled pita with a side of brasillian fries \$12

Figure 2. Sample menu <<https://cafebeaudelaire.com/>>.

Note, however, that in a language test, many of these contextual features contributing to authenticity are somewhat difficult to achieve: setting (a classroom, test center, or computer screen at home), participants (teacher/tester, students), purpose (displaying language ability), act sequence (follow the order of item/task presentation), tone (formal, didactic, indirect), language (face-to-face, learned English), norms (academic), and genre ("test"). The menu, in such a situation, would still be genuine, but the authenticity of the test takers' response would be somewhat in question - genuineness travels well, but authenticity is more difficult to achieve. Of course, the creative tester can help bring some authenticity into the

test by using computer technology to simulate a realistic context. The restaurant example could be simulated as in the following image:



Figure 3. Restaurant <<https://www.touchbistro.com/blog/suggestive-selling-techniques-for-restaurants/>>.

The diners are holding menus and the server is taking their order, whilst in an aviation context delivered by computer, M. Park (2014) reproduced a military airport scene for a test of English for Korean helicopter pilots:

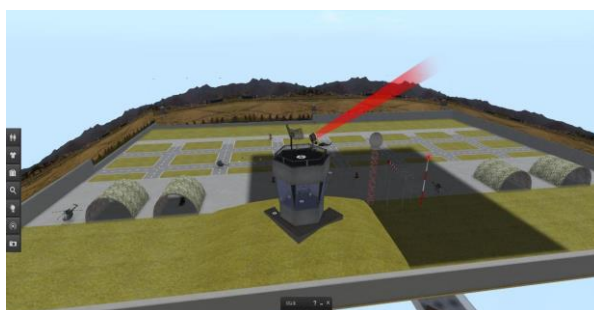


Figure 4. Military airport (M. Park 2014).

Thus, in a specific purpose language test, authenticity is a major factor in helping assure that test takers will respond to the test tasks in a manner interpretable as evidence of their communicative ability in the target language use situation.

### 3. Inseparability

Recall that in the Codes of Relevance model, the language of the innermost professional circles represents the highest levels of specialist background knowledge and language knowledge. The more specific the language, the more background knowledge is required for communication. A problem for specific purpose language testing – and for more general purpose assessment as well – has been the interaction between language knowledge and background knowledge. Clearly, in order to participate in communication about any topic, one must know something about that topic, or at least know enough to ask intelligent questions. In specific purpose language assessment, background knowledge is a part of the test construct (D. Douglas, 2013) and is thus an *inseparable* part of what is being assessed. A number of recent research studies have confirmed this.

Y. Cai and A. Kunnan (2018) examined the inseparability of content knowledge from LSP reading ability by having 1,491 nursing students in China respond to a nursing English test and a nursing knowledge test. Results showed that the domain-specific passage factors were significantly correlated with their corresponding domain-specific nursing knowledge factors, while domain-general nursing knowledge significantly predicted the variance of the domain-general reading factor. They concluded that content knowledge is inseparable from LSP reading ability.

U. Knoch 2014 studied the validity of the ICAO Rating Scale. She had ten pilots listen to performances of test takers taking a variety of aviation English tests. The pilots were asked to rate the acceptability of each test taker's language for (a) communicating with other pilots and (b) radiotelephony communications with air traffic control. The results showed that the pilots focused on some but not all of the criteria on the ICAO scale, but that they also often focused on the speakers' technical knowledge, suggesting that technical knowledge is an important component of specific purpose communication ability.

This is implicitly acknowledged in the International Civil Aviation Organization (ICAO) *Manual on the Implementation of ICAO Language Proficiency Requirements* (ICAO Doc. 9835, 2010):

[Aeronautical radiotelephony communications] ... are highly context-dependent since they rely on a great deal of specific technical knowledge related to aviation themes or topics such as aircraft, navigation, air traffic control procedures and equipment..." (Section 3.3.1.b).

Nevertheless, the same document prohibits the express evaluation of technical knowledge: "Language tests should not assess either operational skills or the specific technical knowledge of operations" (Section 6.3.2.10). But as the research studies above suggest, technical knowledge can support a lower level of language knowledge, and so it may be that this limitation should be reconsidered.

## Conclusion

In conclusion, there is no doubt about the importance of the concepts of specificity, authenticity, and inseparability in ensuring the quality of tests of aviation English. In a survey conducted by the Lancaster Language Testing Research Group (C. Alderson 2010) it was found that among several such tests available, little evidence was available regarding their quality, particularly in terms of the validity of interpretations of test performance. The survey team concluded that they "can have little confidence in the meaningfulness, reliability, and validity of several of the aviation language tests currently available for licensure" (2010: 51) and recommend that some international organization begin to monitor the quality of tests in this critical area of assessment; indeed, the International Language Testing Association has begun consultations with the ICAO for this very purpose. More recently, however, N. Bullock and C. Westbrook (2021) found that there remain serious concerns over the validity, reliability and quality of test instruments and that there is a wide disconnect between those being tested and those providing the tests. J. Read and U. Knoch (2009) and B. Grande et al. (2022) provide overviews of the issues involved in the development and interpretation of tests of Aviation English. As discussed above (U. Knoch 2014), there is also concern for the validity of the ICAO scale itself for rating professional Aviation English.



What does emphasizing specificity, authenticity, and inseparability mean for aviation English test development? Clearly, test input should be set at an appropriate level of relevance, and tasks should reflect contextual features and what test takers need to do in their roles in aviation. These might include the following:

- Ability to tolerate and comprehend different varieties of English  
Task: Listen to and demonstrate comprehension of a variety of English accents, discourse styles
- Ability to negotiate meaning  
Task: In a face-to-face interview, interlocutor produces an intentionally ambiguous utterance, the meaning of which must be negotiated by the candidate
- Ability to use (or adjust) phonological features crucial for intelligibility  
Task: Interlocutor indicates a failure to comprehend candidate's utterance, requesting a recast
- Awareness of appropriate pragmatics  
Task: Discourse completion in an aviation context.

Attending to the three factors of specificity, authenticity, and inseparability will lead to more valid interpretations of aviation English test performances.

## References

- Alderson, C. (2010), *A Survey of Aviation English Tests*, (in:) "Language Testing" 27:1, 51–72.
- Bullock, N. / C. Westbrook (2021), *Testing in ESP: Approaches in aviation and maritime English*, (in:) B. Lanteigne, C. Coombe and J.D. Brown (eds.), "Challenges in language testing around the world". Springer: New York, 67–77.
- Cai, Y. / A. Kunnan (2018), *Examining the inseparability of content knowledge from LSP reading ability: an approach combining bifactor-multidimensional item response theory and structural equation modeling*, (in:) "Language Assessment Quarterly" 15:2, 109–129.
- Douglas, D. (2000), *Assessing Languages for Specific Purposes*. Cambridge University Press: Cambridge.
- Douglas, D. (2013), *ESP and assessment*, (in:) B. Paltridge and S. Starfield (eds.), "The Handbook of English for Specific Purposes". Chichester, UK: Wiley-Blackwell, 367–384.
- Elder, C. (2001). *Assessing the language proficiency of teachers: are there any border controls?*, (in:) "Language Testing", 18.2, 149–170.
- Elder, C. (2015), *Exploring the limits of authenticity in LSP testing: The case of a specific-purpose language test for health professionals*, (in:) "Language Testing" 33.2, 147–152.
- Grande, B./ C. Leishman / H. Skilleås (2022), *"I can't come to the words": Effects of including military flavour when testing the oral proficiency of Norwegian joint terminal attack controllers*, (in:) "English for Specific Purposes" 68, 73–86.
- Hymes, D. (1974), *Foundations in sociolinguistics: An ethnographic approach*. The University of Pennsylvania Press: Philadelphia.
- International Civil Aviation Organization (2010), *Manual on the Implementation of ICAO Language Proficiency Requirements*, ICAO Doc. 9835, Montréal.

- Knoch, U. (2014), *Using subject specialists to validate an ESP rating scale: The case of the International Civil Aviation Organization (ICAO) rating scale*, (in:) "English for Specific Purposes" 33, 77–86.
- Knoch, U. / S. Macqueen (2020), *Assessing English for Professional Purposes*. Routledge: London.
- Park, M. (2014), *Development and Validation of Virtual Interactive Tasks for an Aviation English Assessment*. Doctoral Dissertation, Iowa State University.
- Read, J. / U. Knoch (2009), *Clearing the Air: Applied Linguistic Perspectives on Aviation Communication*, (in:) "Australian Review of Applied Linguistics" 32.2, 1–21.
- Widdowson, H. (1979), *Explorations in Applied Linguistics*. Oxford University Press.