Corpus Pragmatics (2017) 1:297–326
DOI 10.1007/s41701-017-0015-x

ORIGINAL PAPER

Corpus Analysis: Pragmatic Conclusions

Kieran Harrington[1]

Received: 14 December 2016 / Accepted: 14 April 2017 / Published online: 24 April 2017
© Springer International Publishing AG 2017

Abstract This paper derives from a comprehensive study of the communication of a community of speakers of other languages living in an asylum seeker centre.The specific focus is on the corpus-based investigation of 48,000 words of lingua franca English transcribed from recordings of interaction between residents of the centre and native-English speaking staff and amongst the residents themselves over a 3 year period. The study is unique in the sense that the speakers are not EFL students, English is not the lingua franca of choice, and the speakers are 'inmates' in a 'total institution'. The basic core spoken vocabulary of the community was gauged by identifying the point in a computer-based frequency count at which frequency dropped off sharply. The presence of broad categories of basic spoken vocabulary was also calibrated. The results, which bear evidence of a vastly limited language system, especially as far as lexis is concerned, raise questions with regard to the interactional achievement of the asylum seekers in transaction and negotiation with native speakers of English. The answer to these questions was provided by the identification of key words in frequency lists and the analysis of the pragmatic function of these words in the detail of the immediacy of talk-in-interaction.

Keywords Corpus Linguistics · Ethnography · Conversation Analysis

Introduction

This paper is based on a study I carried out into the communication of asylum seekers through the medium of lingua franca English, underpinned chiefly by corpus linguistics, and ethnography (Harrington 2016, 2018). I worked in an asylum seeker

✉ Kieran Harrington
kierharrington@gmail.com

[1] Galway, Ireland

centre[1] as an education coordinator for 3 years and during this time I collected ethnographic data and a corpus of 98,000 words.

As there was no single homogenous culture with its concomitant schematic knowledge, and no common 'foreign' language (there were over 30 nationalities in the centre and over 50 languages) that could be easily and readily accessed with regard to possible sociopragmatic or pragmalinguistic transfer, the ethnography focussed on the common 'culture' of persecution and flight, and the adaptation to the tedium and confinement of institutionalization. The asylum seekers' own 'dialogical' representation of their histories and their institutionalization demonstrated how such experience impacted their ways of speaking.

The 98,000 words were transcribed from recordings of 31 communicative events involving interaction between residents of the centre and the English-speaking staff (including ethnographic interviews), but also interaction amongst the residents themselves. In order to characterize the lingua franca, 48,000 words pertaining to the residents only were isolated from the main corpus and were analysed using typical corpus analysis techniques. The complete corpus was then accessed in order to analyse the use of the lingua franca in talk in interaction, and specifically items highlighted by the corpus analysis. The initial investigatory process was naturalistic in the sense that it moved from the study of people (through ethnography), to the study of the lingua franca (through corpus linguistics) to the study of the use of that language with the English-speaking staff of the centre (using detailed analysis of talk in interaction). But the blend of ethnography, corpus linguistics and conversation analysis was also cyclical, as the language analysis, consisting of detailed examination of talk-in-interaction, benefitted from constant back and forth referral to the ethnography and corpus analysis. This robust methodological blend is represented in Fig. 1.

This paper focusses on the lingua franca and its use, as elucidated by the interplay between corpus analysis and detailed analysis of talk-in-interaction. While most studies on English as a lingua franca seem to prioritize transient or virtual communities of students and academics (see for example, Jenkins 2000, 2009; Prodromou 2008; Seidlhofer 2010; Cogo and Dewey 2012) with high levels of English, this study focusses on a closed community of speakers of other languages who used English to negotiate the daily ordeal of living in a reception centre. Their levels ranged from those with only a few words to those, as in the case of Nigerians, who spoke it as a second language.

In "Vocabulary and Basic Components of Spoken Conversation" section, the lingua franca is characterized following the methodology and parameters of McCarthy (1999), when he considered the question of what constitutes a basic vocabulary for spoken communication. This is followed in "Word Frequency" section by a closer analysis of word frequency in the *lingua franca* corpus and the identification of key items, which in the final section are qualitatively examined in the context of talk-in-interaction ("Minimal Responses in Talk-in Interaction" section).

[1] All geographical and personal names have been anonymized.

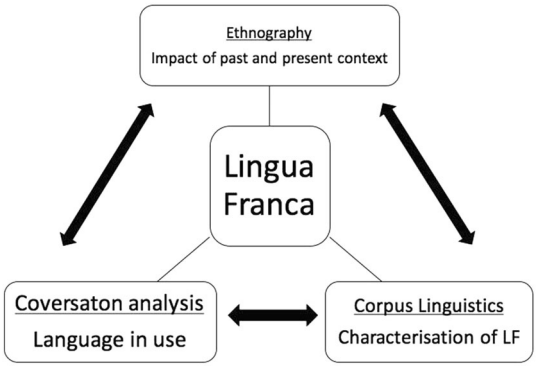

Fig. 1 The synergistic blend of ethnography, corpus analysis and conversation analysis

Vocabulary and Basic Components of Spoken Conversation

Vocabulary Size and Coverage

McCarthy (1999) based his study on a 3-million-word sample of the 5-million-word Cambridge and Nottingham Corpus of Discourse in English (CANCODE) transcribed from conversations of native speakers in settings such as shops, offices and private homes. McCarthy (1999: 236) considered that vocabulary coverage can be gauged in computer-based frequency counts 'where frequency drops off rather sharply, from hard-working words which are of extremely high frequency to words that occur relatively infrequently, in other words, the frequencies do not decline at a regular rate…' The point at which this drop is discernible ('with a marked decrease in the number of words that occur more than 100 times') in McCarthy's study is identified after the first 2000 words and this is seen as a boundary between the core and the rest and wherein lies the 'heavy-duty core vocabulary' [see also Carter (1998) on core vocabulary, Nation (2006) on vocabulary needed for reading and listening, Van Zeeland and Schmitt (2012) for listening comprehension, and Schmitt et al. (2015) for review of vocabulary coverage].

In the lingua franca corpus (LFC henceforth) of the residents of the centre (see Fig. 2), the first 300 words do most of the work and constitute 80% of all the vocabulary. The word 'work' here, of course, refers to their daily negotiating and transacting functions in their pursuit of basic supplies and services in the asylum seeker centre, and in their casual interactions with one another. Even more noteworthy is the fact that the first 100 words account for 64% of the total words—and this is where a sudden drop appears. The residents of the centre, then, interact, transact and negotiate chiefly with a frequent vocabulary of 100 words. This constitutes the primary indication of the limitations of the language system.

Nine Broad Categories of Basic Spoken Vocabulary

McCarthy (1999: 240) identified in the first 2000 words of the 3-million-word CANCODE corpus 'fairly clear categories' of language use which constituted

100	200	300	400	500	600	700	800	900
64%	10.4%	5.25%	3.56%	2.30%	2%	1.60%	1%	0.94%

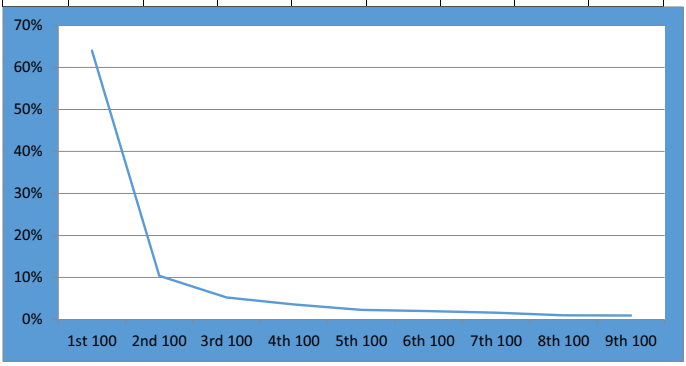

Fig. 2 Word coverage and frequency bands for every hundred words

'important components of basic communication.' Here I will gauge the LFC for the frequency of the components of these 'nine broad categories of basic spoken vocabulary'. They are as follows:

- Modals
- Delexical verbs
- Interactive words
- Discourse markers
- Basic nouns
- General deictics
- Basic adjectives
- Basic adverbs
- Basic verbs for actions and events

Modals

Modal items refer to degree of certainty (epistemic modality) or necessity (deontic modality) and include (in McCarthy's categorization of 'high frequency items'), the non-lexical modal verbs *can, could, may, might, must, should,* the lexical verbs, *look, seem* and *sound,* the adjectives *possible* and *certain* and the adverbs, *possibly*, *certainly* and *apparently* (see Carter and McCarthy 2006 for extended list). In the LFC *can* is the only one used with any degree of frequency. It figures in the first 100 most frequent words in the corpus (38th) which represents 5298 uses when normalized to one million, and is used in 23 of the 31 corpus texts (Table 1).

Table 1 Modal items

Rank	Modal	Freq.	Per M	Range
38	Can	206	5298	23
222	Could	23	592	7
224	Must	23	592	8
247	Possible	20	515	8
295	Should	17	437	10
397	Probably	12	308	6
422	Might	11	282	4
479	May	9	231	5
932	Seem	4	103	3
1827	Certain	1	25	1

Range refers to the 31 communicative texts that were recorded and transcribed, that is, to the number of texts that the word was used in. While this is referred to minimally below, it is useful for information purposes as it also shows how widespread usage of a particular word is

Table 2 Delexical verbs

Rank	Verb	Freq.	Per M
150	Get	39	1005
220	Make	23	593
235	Take	21	540
379	Do	12	309

Delexical Verbs

This category includes verbs such as *do, make, take* and *get* when they are not used in their lexical sense, but rather when used collocating with nouns, prepositional phrases and particles to make meaning. As Table 2 shows, these delexical verbs are not frequent in the LFC, indeed none of them are in the first one hundred words.

Interactive Words

Interactive words, also known as 'stance' words (see O'Keeffe et al. 2007), express the speaker's attitude in the communicative situation and are useful in softening face-threatening utterances. Given the relative paucity of these items in the LFC (see Table 3), the following assessment by McCarthy (1999: 242) of the consequences of their absence becomes even more significant:

> These are absolutely central to communicative well-being, to creating and maintaining appropriate social relations. They are therefore not a luxury, and it is hard to conceive of anything but the most sterile and banal survival-level communication occurring without their frequent use. The speaker who cannot use them is an impoverished speaker, from an interpersonal viewpoint.

The frequencies of the words *just, thing(s) whatever, anyway, basically, a bit, really, quite,* and *literally* are shown in Table 3. While the word *just* is within the first 100 words of the LFC, and is distributed across 20 of the 31 texts, its per million value is less than half that of the same item in CANCODE.

Table 3 Interactive words

Rank	Item	Freq.	Per M	Range
55	Just	121	3112	20
116	Really	52	1337	12
163	Thing (s)	35	900	10
706	Whatever	6	154	4
984	Anyway	3	77	3
0	Basically	0	0	0
0	A bit	0	0	0
0	Quite	0	0	0
0	Literally	0	0	0

Table 4 Discourse markers

Rank	Item	Freq.	Range
34	You know	235	22
54	So	125	21
409	I mean	11	7
439	Right	10	7
457	Well	9	6
984	Anyway	3	1

Discourse Markers

An early reference to discourse markers (to *well* specifically) by Labov and Fanshel (1977: 156) described their function of referring back to 'some topic that is already shared knowledge among participants,' while Schiffrin (1987: 31), examining 12 discourse markers, defines them as 'sequentially dependent elements which bracket units of talk,' In the study under comparison here, McCarthy (1999: 243) points to the frequent use of *you know*, *I mean, right, well, so, good* and *anyway* as discourse markers which 'organise the talk and monitor its progress'. Their absence, he concludes, leaves conversational participants 'disempowered and at the risk of becoming a second class participant.'

Only *you know* and *so* figure with any notable frequency in the LFC (Table 4)—within the first 50 most frequent items; however, on closer analysis it was found that usage was limited mainly to one-to-one ethnographic interviews and rare in the discourse of daily interaction with the English-speaking staff.

Canonical discourse markers such as *right* (30th most frequent item in CANCODE) and *well* (27th most frequent in CANCODE) are noticeable for their infrequency. *Well* is only used nine times (457th in the word frequency list), and *right* ten times (439th) discoursally, although, as can be seen from the ten instances of *right* (Fig. 3), its function is ambiguous, and can only be said to function as an indicator of new action and sequence (Carter and McCarthy 2006: 136) in examples 3,4, 6 and 10. For this function, as will be seen in the "Minimal Responses in Talk-in Interaction" section, the residents frequently resort to *yeah*.

1. Native speaker: Just a second I'll see who is at the door.
   Said (Moroccan male): **Right**

2. Holly (Nigerian female): Sit on the table
   Boris (Georgian male): Er **right**.

3. Miro (Serbian male): **Right** ladies first

4. Oleg (Ukrainian male): Integrate I talk about this need tall sometimes in English **right** help one another

5. Princess (Nigerian female): Some too are Christians **right**.

6. Princess: Yeah so there is no need to. **right** meeting people

7. Brane: Er name of house is St Bridget's.
   Holly: Oh **right**

8. NS: Sometimes you see the ads like that apartment above shops
   Anna (Belorussian female): **Right**

9. Boris: Yep. An ah not. r**igh**t, ah. from our open at one corner

Fig. 3 *Right* as a discourse marker

General Deictics

Lyons (1971: 275) speaks of 'deictic categories' which 'handle the "orientation" features of language which are relative to the time and space of utterance.' Deictic terms situate the speaker in the world in terms of particular times and places, and they cannot be understood without such context. They include the demonstratives, *this, that, those, these,* the locational adverbs *here, there,* and the temporal adverbs *now, then* and *ago*. These all appear within the first 500 words of the top 2000 word forms identified by McCarthy (1999), as pointed out by O'Keeffe, McCarthy and Carter (2007: 37). Table 5 shows that such items are frequent in the LFC; *this, that, now* and *here* all figure in the first 100 most frequent items.

Basic Nouns

McCarthy (1999) provides a long list of generic nouns which are frequent in basic core vocabulary. Table 6 shows the entries for such basic nouns in the first one hundred words of the LFC, in which there are eight generic nouns.

The high frequency of *people, language* and *country* and their position within the first one hundred words is not surprising, given that conversation in the centre

Table 5 Deictics

Rank	Deictic	Freq.	Per M	Range
32	This	242	6224	25
35	That	210	5401	25
62	Now	117	3009	22
64	Here	116	2984	20
106	Then	57	1466	17
131	There	47	1209	14
173	Ago	31	797	10
172	Back	32	797	10
254	Those	19	489	8
298	These	16	412	7

Table 6 Basic nouns

Rank	Noun	Freq.	Per M.	Range
34	People	230	5967	21
59	Language	119	3065	17
73	Country	99	2550	19
75	Time	89	2289	23
76	Year(s)	88	2263	21
84	Problem	74	1906	17
96	School	65	1672	14
99	House	61	1571	11

frequently revolved around issues of nationality, language and culture. The following examples (Fig. 4) from the corpus (mostly from ethnographic interviews and not daily interaction) show how these terms occur in the context of culture, religion, languages and nationality.

Other basic nouns (Table 7) mentioned by McCarthy—*life, children, kids,* and *car*—could not be classified as frequent in the LFC. None appear in the first 100 words, and only *life* appears in the second 100 (101–200).

Basic Adjectives

McCarthy (1999) identifies the adjectives *lovely, nice, good, bad, horrible, terrible,* and *different,* which communicate evaluations of people, situations, events and things, as occurring in a frequent manner, within the first 2000 core words of vocabulary. Table 8 shows that only *good* and *different* occur within the first one hundred words of the LFC frequency list.

Of the 144 uses of *good,* only 30 (21%) are used attributively. The remainder are used principally in the utterance 'That's good' or It's good', the lowest level users of English frequently omitting the subject pronouns, as in the following examples from the corpus:

People

I not pray with the African people
My idea Irish people are good
People from Congo they are very kind people
How happened er that two people culture joined in on state
Islam yes so people are Muslims
The most people live in Somalia

Language

Tshiluba is the language of my family
You know same language Albanians
My language Oromo
Urdu is national language
Before it was Serb and Croat language
English is the official language

Country

Sudan like other country like Algeria
Our country is cut off outside the culture
Yes in my country before this war
The time when drought have my country

Fig. 4 *People, language* and *country*

Table 7 Other basic nouns

Rank	Noun	Freq.	Per M.	Range
189	Life	28	720	9
302	Children	16	412	8
334	Car	14	360	6
661	Kids	6	154	4

Table 8 Basic adjectives

Rank	Adjective	Freq.	Per M	Range
48	Good	144	3711	22
85	Different	73	2527	14
151	Bad	38	979	11
218	Nice	24	619	14
1727	Awful	1	26	1
2380	Lovely	1	26	1
0	Terrible	0	0	0

Registration interviewer (female): Sometimes there is snow.
Princess (female speaker of Nigerian Pidgin English): Yeah is **good**.

Anyeta (Ukrainian female): They can be like these small kids.
Katya (Ukrainian female): Yeah is **good**

NS Male: What that language
Said (Moroccan male): Yeah is good

Oleg (Ukrainian male): Not too fast er correct correct is **good.**

Fig. 5 *Good*

The frequency of the adjective *different* is, as in the case of the nouns *language*, *country* and *people*, not surprising, given the preponderance of issues of nationality, language and culture in the conversation of the residents, especially in the ethnographic interviews. In fact, all of the 73 uses of *different* in the corpus refer to these three topics, the adjective specifically collocating with the nouns *language* 20 times, with *country* 12 times, and with *people* 10 times (Fig. 5).

Basic Adverbs

McCarthy (1999) classifies as extremely frequent adverbs which refer to time—*today, yesterday, tomorrow*, *eventually, finally*, to frequency—*usually, normally, generally*, manner—*suddenly*, *fast, totally, especially,* to stance—*basically*, *hopefully*, *personally* and *literally,* and prepositional phrase adverbials—*in the end* and *at the moment*. The frequency, and distribution across texts, of these adverbs in the LFC (represented in Table 9) is extremely low.

The most frequent adverbs are, with the exception of *especially,* the specific time referents, *tomorrow* and *today,* most of which concern appointments. In fact, as seen in Fig. 6, all 18 uses of *tomorrow* refer to appointments and arrangements.

Basic Verbs for Actions and Events

McCarthy lists the verbs *sit, give, say, leave, stop, help, feel, put, listen, explain, love, eat,* and *enjoy* as forming part of the basic vocabulary for expressing everyday activity. Their lemmatized frequency is shown in Table 10. Only *say* figures in the first one hundred most frequent words, but it is used in only 18 of the 31 texts.

Worthy of note is that verbs that refer to communication are much more frequent than reporting verbs in the LFC (Table 11), the reverse of the normal large-corpus pattern. Of the communication verbs, *talk* (lemmatized) is 44th in the list of most frequent words, and far exceeds frequency rates in larger corpora. *Speak* follows a similar pattern.

One might conclude that the residents use the verbs *talk* and *speak* for both the reference to an act of communication and as reporting verb; however, a manual

Table 9 Basic adverbs

Rank	Adverb	Freq.	Per M	Range
270	Especially	18	464	8
282	Tomorrow	18	464	10
403	Today	12	309	7
746	Fast	5	128	4
959	Usually	4	103	2
966	Yesterday	4	103	3
1113	Normally	3	0	2
0	Eventually	0	0	0
0	Finally	0	0	0
0	Basically	0	0	0
0	Generally	0	0	0
0	Suddenly	0	0	0
0	Hopefully	0	0	0
0	Personally	0	0	0
0	Literally	0	0	0
0	At the moment	0	0	0
0	In the end	0	0	0

1. I ask my friend **tomorrow**
2. **Tomorrow** I need to go to town
3. I'm not **tomorrow**
4. I come **tomorrow**
5. Alright so I see then **tomorrow**
6. Manhana **tomorrow** ten o'clock
7. I think about being alive **tomorrow**
8. So today and **tomorrow**
9. Today Friday **tomorrow** Saturday
10. Yes ah **tomorrow** afternoon
11. I try to eat less **tomorrow**
12. When **tomorrow**
13. Is it convenient for you **tomorrow**
14. I see then **tomorrow**
15. What about **tomorrow**
16. You can come **tomorrow**
17. See you **tomorrow**
18. Anne Marie **tomorrow**

Fig. 6 *Tomorrow*

Table 10 Everyday lexical verbs

Rank	Verb	Freq.	Per M	Range
63	say	116	2983	18
178	eat	30	771	9
216	Help	24	617	11
244	Listen	21	540	11
307	Feel	16	412	6
312	Put	16	412	8
375	Sit	13	334	7
436	Stop	11	282	4
445	Leave	10	257	5
670	Love	6	153	4

Table 11 Verbs of communication and reporting

Rank	Verb	Freq.	Per M	Range
44	Talk	171	4398	18
49	Speak	133	3420	24
63	Say	116	2983	18
99	Tell	61	1568	16
	Total	479	12,369	31

search of concordance lines did not uncover any example of *speak* being used as a (deviant) reporting verb and only uncovered the following two examples of *talk* employed in this way.

'I go with my leg and I talk Eugene go back' (Oleg, Ukrainian male)
'I talk okay where are you?' (Boris, Georgian male)

The manual search also revealed that the residents simply avoided indirect reporting, and this ties with an overall pattern of avoidance, simplification, and exploitation *ad maximum* of certain words (such as *just, good, speak, talk,* and *can)*. The ethnography in the larger study showed that the residents of the centre communicated mainly for the purposes of survival—there was little casual interaction. The infrequency of such basic and familiar verb as *sit* and *stop* then points more to economization of communication rather than, perhaps, lack of knowledge. The picture that the survey of word coverage and the nine broad categories of spoken vocabulary paints is of an extremely reduced language system. The question that arises then is: how did the residents communicate with the English-speaking staff on a daily basis with such a reduced vocabulary at their disposal, especially in the more complicated communicative situations such as those of requesting finance and funding? Examination of word frequency lists will provide further clues.

Word Frequency

In this section the word frequency list of the LFC, the word frequency list of a specific negotiating interaction which took place in the staff offices, which I refer to as the 'Funding Corpus' and the vocabulary usage of the 'negotiators'—residents who had to attend for meetings in staff offices to requesting finance and funding—will be examined. I begin with the complete LFC in "Word Frequency List of the LFC" section

Word Frequency List of the LFC

Table 12 shows the 50 most frequent items in the LFC normalized to one million words. CANCODE (C) is used as the comparative corpus to maintain coherence

Table 12 LFC word frequency list

R	Word	Per M	C	Per M	R	Word	Per M	C	Per M
1	I	38,500	The	33,867	26	Are	7900	Like	6787
2	Yeah	31,100	I	30,197	27	Because	7200	Well	6786
3	Er	30,600	And	28,242	28	One	6700	What	6641
4	You	28,200	You	27,504	29	My	6600	Do	6574
5	And	22,300	It	21,249	30	Okay	6600	Right	6310
6	To	21,800	To	21,170	31	Was	6500	Just	6237
7	The	20,800	A	20,704	32	This	6200	He	6135
8	No	18,700	Yeah	18,296	33	So	6200	For	5969
9	Is	16,800	That	16,986	34	People	5900	Erm	5688
10	In	16,400	Of	15,641	35	From	5400	This	5626
11	They	12,400	In	12,559	36	What	5300	Be	5617
12	Know	12,200	Was	10,080	37	Me	5300	All	5536
13	Yes	10,500	it's	9567	38	Can	5200	There	5255
14	But	10,400	know	9329	39	Very	5200	Got	5226
15	A	9900	Is	9089	40	If	5000	that's	5138
16	We	9800	Mm	8820	41	About	4600	Not	5094
17	Not	9200	Er	8695	42	There	4500	don't	5041
18	For	8900	But	8306	43	Do	4500	If	4886
19	Don't	8900	So	8014	44	He	4300	Think	4860
20	Have	8900	They	7772	45	Ah	4100	One	4778
21	That	8600	On	7140	46	Go	4100	With	4575
22	Like	8500	Have	7123	47	Good	4000	At	4438
23	It's	8300	We	7117	48	With	3700	Or	4283
24	It	8300	Oh	7045	49	Understand	3400	Then	4240
25	Of	8200	No	7017	50	Them	3300	She	4123

Values are normalized to per million frequencies. CANCODE (C) frequencies (columns 4 and 9) are included for comparison. R refers to rank of frequency

with the preceding analysis of coverage and the nine broad categories of spoken vocabulary. There are multiple distinctive features that could be commented on, such as the reduced use of the definite article, 1th in CANCODE and all other large corpora, but 7th in the LFC, due perhaps to both the paucity of common nouns in the corpus and to the avoidance of articles, especially by residents who speak Slavonic languages. The pause filler *er,* which I have included as both *er* and *erm* figure in the CANCODE list, is extremely frequent, due largely to the disfluency that might be expected in the discourse of people speaking in a foreign language, especially those at lower levels. It appears more frequently in more expansive turns (see "Funding Corpus Word Frequency Lists" section), functioning to hold or fill as the speaker performs lexical or morphosyntactic searching; conversely, it scarcely occurs in shorter turns where the residents use more motivated minimal responses, such as *yeah*. This minimal response is the second most frequent item in the corpus and is used much more frequently than in CANCODE—that is 31,000 per million (LFC to 18,296 per million CANCODE), taking on multiple functions, as will be seen below, such as fulfilling the discoursal role normally associated with the high-frequency discourse markers *right* and *well* (McCarthy 1999; O'Keeffe et al. 2007) which only occur ten and nine times respectively in the LFC corpus.

Funding Corpus Word Frequency Lists

The 'Funding Corpus' of 5000 words, a sub-corpus of the LFC, was gathered from a recording of visits by the residents to a staff office to request a form to apply for financial assistance that was available for sporting activities, to request information on the process, to explain individual situations and to request finance for other related expenses. The written transcriptions of the longer stretches of discourse or longer turns of the residents tended towards disfluency and incoherence, with a use of language that could be called sub-standard English, as in the following example:

> Er my daughter every very Saturday go taxi way city and go for bus one euro … and go and sometimes one mother er. is friend and wait where is Tesco and .. bring you where is and every time I must ask please do you want take my daughter and er er …bus for one euro it's er. too much late and knows er every Fri go and small time play

However, those of us who carried out these funding 'interviews' came away with the impression that the English of the 'negotiators' was quite good and fluent, and, there was certainly no negative effect on the funding applications, as most of the people who applied for financial help were successful and even received extra amounts for other financial concerns which were expressed at these meetings.

The word frequency list of the Funding Corpus (Table 13) contains clues as to how an impression of fluency (rather than accuracy) is created. Immediately striking is the high frequency of the minimal response *yeah,* replicating the pattern evidenced in the larger corpus. It accounts for 6.7% of the vocabulary items. Furthermore, if the occurrences of the other minimal responses are added together, the aggregate percentage is 12.8, which gives a clear idea of the dependency on these items.

Table 13 Word frequency list of Funding Corpus

Rank	Item	F	%
1	Yeah	93	6.7
2	I	74	5.3
3	And	47	3.4
4	Okay	46	3.3
5	You	46	3.3
6	Er	32	2.1
7	For	25	1.8
8	Know	25	1.8
9	In	20	1.5
10	Is	20	1.5
11	My	20	1.5
12	But	19	1.4
13	No	19	1.4
14	Yes	19	1.4
15	The	18	1.3
16	Like	16	1.2
17	To	16	1.2
18	Go	15	1.2
19	Me	13	1.1
20	A	12	1
21	Want	12	1
22	Can	11	1
23	IT'S	11	1
24	Not	11	0.9
25	One	11	0.8

The word frequency lists of the individual negotiators follow similar patterns. As can be seen in Table 14, the minimal response *yeah* accounts for 12.98% of the discourse of Dragana (a female Serbia resident), while aggregated frequencies for *yeah, okay* and *yes* amount to 20.5% of her total output. One would expect the petitioner of funds to take the onus, as it is she who is requesting the funding, but the dependency on minimal responses, associated with turn yielding in listenership (see Fries 1952; Kendon 1967; Gardner 2001; O'Keeffe and Adolphs 2008), seems to suggest submissiveness. Field notes taken at the time of the interaction, however, confirm that the funder in question came away with the impression that Dragana had lead the conversation and that her English was of a high level. She also managed to negotiate a tacit agreement from the funder to fund a course (which had already started) for her daughter, a taxi to get her to the sports complex and additional funding for another daughter and her husband.

While the minimal response *yeah* is not as pervasive in the discourse of Darya (a Ukrainian female), it still figures as the second most used item, and *okay*, also used as a minimal response is also frequent (Table 15).

Table 14 Dragana

Rank	Item	F	%	Rank	Item	F	%
1	Yeah	57	12.98	11	Go	9	2.05
2	And	26	5.92	12	Know	9	2.05
3	Er	22	5.01	13	Daughter	8	1.82
4	I	21	4.78	14	Time	8	1.82
5	Okay	19	4.33	15	Is	6	1.37
6	My	15	3.42	16	Four	5	1.14
7	Yes	14	3.19	17	It's	5	1.14
8	For	13	2.96	18	No	5	1.14
9	You	11	2.51	19	One	5	1.14
10	Every	9	2.05	20	Play	5	1.14

Table 15 Darya (Ukraine)

Rank	Item	F	%	Rank	Item	F	%
1	I	30	7.81	11	Okay	7	1.82
2	Yeah	15	3.91	12	She	7	1.82
3	In	14	3.65	13	Like	6	1.56
4	And	12	3.13	14	Er	5	1.3
5	The	11	2.86	15	Is	5	1.3
6	Know	10	2.6	16	Oh	5	1.3
7	No	9	2.34	17	One	5	1.3
8	But	8	2.08	18	This	5	1.3
9	Me	8	2.08	19	Can	4	1.04
10	You	8	2.08	20	Fitness	4	1.04

The next word frequency list (Table 16) corresponds to a meeting between the funder and Anyeta, a Ukrainian female, and Katya, another Ukrainian female, who has come as a translator—which she announces at the beginning of the meeting—to help her compatriot to request funding for her teenage son. Again, *yeah* is the most frequent item, accounting for 5.92% of the total output.

The preponderance of minimal responses and particularly *yeah* in the Funding Corpus suggests that these items are vital in the interaction. In the next section I present some examples of the use of minimal responses within the here and now of interaction to demonstrate their general impact and their specific multifunctionality, which in some cases moves beyond previously reported uses of these features in the literature which I briefly review below.

Minimal Responses in Talk-in Interaction

Minimal Responses in the Literature

In this paper, hitherto, I have used the term *minimal response* for discourse features such as *yeah* and *okay*, *uh huh* and *mm*. However, such features have been termed in

Table 16 Anyeta and Katya (Ukraine)

Rank	Item	F	%	Rank	Item	F	%
1	Yeah	21	5.92	11	Not	7	1.97
2	You	17	4.79	12	That	7	1.97
3	I	14	3.94	13	Because	5	1.41
4	Like	9	2.54	14	Can	5	1.41
5	They	9	2.54	15	Find	5	1.41
6	For	8	2.25	16	It	5	1.41
7	To	8	2.25	17	Will	5	1.41
8	And	7	1.97	18	Go	4	1.13
9	But	7	1.97	19	Have	4	1.13
10	Is	7	1.97	20	Know	4	1.13

multiple ways beginning with their first mention by Fries (1952) when he referred to them as 'signals of continued attention'. There is not sufficient space here to cover the full gamut of appellations, but some significant examples include 'accompaniment signals' (Kendon 1967), 'verbal listener responses' (Dittman and Llewellyn 1968), 'backchannel responses' (Yngve 1970), 'assent terms' (Schegloff 1972), 'acknowledgement tokens' (Sacks et al. 1974.), 'continuer' (Schegloff 1982), 'receipt tokens', (Heritage 1984; Brazil 1995), and 'hearer signals' (Bublitz 1988). The term 'response tokens' (Gardner 1998, 2001), is now widely used, especially by corpus linguists such as Carter and McCarthy (2006), O'Keeffe et al. (2007) and O'Keeffe and Adolphs (2008).

Multiple functions are described for these features of discourse and Schegloff (1982), for example, reflecting on how they seemed to be viewed as the detritus of interaction, points out that they are vital to successful interaction. Worthy of note, especially in the context of the present paper, is Fellegy's (1995: 196) consideration that the 'minimal response is a hinge between the linguistic and communicative competence of listeners.' Gardner (2001: 1–14) summarises their 'exquisite complexity' and multifunctionality. For him, they link discourse, project a certain course for the ensuing talk, contribute to the management of the turn-taking system, express positive or negative effect, judgement or attitude, acknowledge a prior utterance indicating non-uptake of an opportunity to talk, and mark a heard utterance as news. Stubbe (1998: 258) presents a 'feedback continuum' which reflects the degree of engagement with the speaker's talk, while Jefferson (1993) highlights movement between *uh* and *yeah* toward preparedness for speakership, and Gardner (1998) presents core 'canonical' and 'non-canonical' uses of *mm hm, yeah*. The non-canonical usage of *mm hm, yeah,* and *mm*, pertains when the items take on different intonation contours: he considers that *mm* with falling intonation functions as 'a more retrospective acknowledging token' (1998: 216), while *yeah* and *mm* with a fall-rising tone, rather than functioning as simple continuers, serve to request extra information.

Within corpus linguistics studies, O'Keeffe et al. (2007) and O'Keeffe and Adolphs (2008) distinguish between convergence and engagement tokens. The

former occur where speakers and listeners agree and 'converge on opinions or mundane topics' (O'Keeffe et al. 2007: 150) and a change of topic is collaboratively negotiated. Engagement tokens signal the listener's interest in the ongoing turn and on the affective level demonstrate emotions such as empathy and sympathy—these tokens, however, are mostly non-minimal, the authors including adjectives such as *brilliant*, adverbs such as *absolutely* and phrases such as *you're not serious*?

The Use of Minimal Responses in the Funding Corpus

The first group of examples of the use of minimal pauses example come from a 20 min meeting between Darya and the funder. Darya uses *yeah*, *yes, oh, ah, okay* and *uh huh* with multiple functionality, demonstrated in the first extract below, where 'yeah' is used for compliance (line 2) as she hands over her residence card, as a discourse marker much in the same way as *right,* and as a tag question (Carter and McCarthy 2006: 198) with rising intonation. 'Oh' (line 8) is then used to acknowledge the new information (see Heritage 1984 for *oh* as 'change of state' token) but without the rising intonation associated with this marker in such situations, which might be indicative of her knowing already that swimming was not the only sport available.[2]

Extract 1

1.		**Funder**: Hello come in <10 sec> have a seat there <P45> Can I have a look at it
2.	➔	**Darya** Yeah (1) <10 sec>
3.		**Funder** Have you children
4.		**Darya**: Yes one son one year three months
5.		**Funder**: Okay <P5> okay tell me what you want to do do you understand so what it is
6.	➔	**Darya**: Yeah (2) this is for swimming ↑yeah (3)
7.		**Funder**: No it's for swimming or any sport
8.	➔	**Darya:** Oh -

The funder proceeds to explain the process to her, and she uses 'uh huh' (line 12 and 15) as continuers. The funder then tells her to sign the application form and look for the cheapest places. She responds with 'oh' (line 19, in the middle of his turn), as if she had heard new information. He then continues ('stuff like that…') and asks almost as a tag added to the end of his turn, 'do you understand?' She takes it as a tag, ignoring the question and tells him what she wants (line 21): 'Oh me better I want like er shaping you know.' The 'oh' functions almost as a synonym of 'but' and the 'er' seems to function here as a hedge, rather than the typical pause filler' before she says the word 'shaping'; the marker 'you know' with rising intonation also mitigates the request at the end of her turn, although the prominence and loudness given to the word 'know' serves to reinforce the request. The first requestive hurdle is over—Darya has got the funder to consider 'shaping' rather than swimming as he signals agreement with just one word, 'Okay' (line 24), to which Darya responds 'Better', signaling preference. The funder asks again if she prefers fitness to swimming and she pauses and then confirms 'yeah' (line 27) but mitigates the 'imposition' by using a low voice.

[2] See "Appendix" for transcription system.

Extract 2.

9. **Funder:** I'll I'll explain it to you we've we've received I have received funding so that
10. people here in Dun Cluain can do sports a certain amount of money so to be fair
11. everybody takes away one of these
12. ➔ **Darya**: uh huh
13. **Funder**: signs it and tries to find something in Xxxxxx city
14. that they would be interested in
15. ➔ **Darya**: uh huh
16.
17. **Funder**: okay like swimming or whatever and then they sign it and
18. bring it back to me and er look for the cheapest place <unintelligible>
19. ➔ **Darya:** oh
20. **Funder** stuff like that .do you understand
21. ➔ **Darya**: Oh me better I want like er shaping you know↑
22. **Funder**: Fitness is it
23. **Darya**: Fitness
24. **Funder**: Okay
25. **Darya:** Better
26. **Funder:** Alright er you prefer that to swimming
27. **Darya**: .◦Yeah ◦(4)

The funder then puts a hypothetical question to Darya, 'If you if you if you went swimming would you take your baby with you' (Extract 3. below, line 28) to which Darya answers almost inaudibly,'I can't no.' The funder seems not to have heard and takes her answer as affirmative, responding that he might be able to find her a mother and baby swimming scheme, but Darya responds with continuers, low 'mm' (31) and even-pitch 'yeah' (33), not showing much interest. When the funder senses that there is no answer coming (line 34) after his turn ends, he comes in again and says, 'so whatever you prefer.' Encouraged or nominated to speak, she clarifies:

'Mm… it's not better together but if not possible without him I can okay.' (line 35)

Darya has achieved something here that even a native speaker would have found difficult—that is to downplay the fact that she does not really want to take her baby with her. She seems clear in her mind that she wants to do 'fitness' alone but does not want to come out straight and say it. She delays (or conceals) the expression of her real intentions by using the continuers to simulate interest in the funder's talk about finding a mother and baby course. Then it is the pause line (35) and ambiguity that gets the message across to the funder, who then requests clarification. She states clearly now (line 38) that she would prefer to do the fitness course, and tries to end the topic with a firm 'Thank you, no' (line 40).

Extract 3

28. **Funder**: And. if you if you if you went swimming would you take your baby with you?
29. **Darya**: °I can't no
30. **Funder**: You might I might be able to find a place for babies like er Katya has a baby
31.➔ **Darya** °mm
32. **Funder**: there are a few more girls with babies that you might all be able to go together
33.➔ **Darya** ↔ yeah (5)
34. **Funder**: or with the babies … so whatever you prefer
35. **Darya**: … <syll> mm it's not better together but if not possible without him I can okay
36. **Funder**: <types 3 secs> and which would you prefer the swimming you prefer which would
37. you prefer the fitness or the swimming
38. **Darya**: Better if fitness
39. **Funder**: Anything else
40. **Darya**: Thank you no

A similar pattern continues, Darya now using *yeah*, *ah, okay* and *uh huh* to support and encourage (or hasten) the funder's long explanation of how to access cheap

sports clubs in *****, until she confirms for him with 'Okay.yeah' (line 89) that she will be going by day. She then (Extract 4) begins to seize a little power or equality and the interaction turns into a 'to and fro' dynamic between her and the funder, both of them supporting each other, the response tokens facilitating a fluid discourse. Darya takes a complete turn (line 91), prefacing with 'Okay', to say she will need to ask in the club. She confirms that she will bring the application form back the next day. The funder takes this cue and rephrases and says,'Yeah you bring that back', and she responds 'Yeah' (line 93) for him to continue. The talk then moves toward two separate 'sections', the first (lines 94–102) where Darya supports (or hastens) the funder's general topic with the continuers 'ah' (line 95), 'yeah' (line 97), 'uh huh' (line 99) and 'yeah' (line 101). The second 'section' (103–109) begins when Darya takes the initiative (lines 103–109) as the funder is about to continue: she uses 'okay' (103) to signal that she is taking the turn. She then reverts to 'her' topic of talking to her friend the next day. A total shift occurs here as it is now the funder who supports her with the response tokens 'yeah' (104), 'okay' (line 106) and 'okay' (line 108). This can be considered a significant interactional achievement on the part of Darya: not only is the funder supporting her talk, he is manifesting his agreement with her posture. The fluency of the interaction between lines 94–102 when she supports him and the fluency of lines 103–109 when he supports her is evident. The closing by her (okay↑ on line 109 with rising intonation and prominence of the last syllable) demonstrates her new control of the trajectory of talk.

Extract 4

90. **Funder**: If it's during the day it's no problem because you can go on the bus
91. **Darya:** Okay but I need ask tomorrow I can bring to you this
92. **Funder** Yeah you bring that back to me
93. ➔ **Darya** yeah (8)
94. **Funder** so I get a pile of these and I decide which one
95. ➔ **Darya** Ah
96. **Funder**: Okay
97. ➔ **Darya:** Yeah (9)
98. **Funder:** So I make I have a certain amount of money maybe about fifteen people
99. ➔ **Darya**: uh huh
100. **Funder**: So if you do it fast
101. ➔ **Darya** yeah (10)
102. **Funder**: and a
103. **Darya:** okay because tomorrow I need going to town <syll>
104. **Funder**: Yeah
105. **Darya**: and I talk with this <girl>
106. **Funder**: Okay
107. **Darya**: And she tell me price for the <syll>
108. **Funder**: Okay
109. ➔ **Darya**: And I am back at one thirty . okay↑

Darya's interactional achievement here is further evidenced by the funder then changing the topic (line 110 in Extract 5. below) from the formal financial issue to the more informal social topic of how long she had been in Ireland—an indication also that he is impressed by her English, further demonstrated by his surprise (line 116) that she had learned English in Ireland and not previously in her home country.

Extract 5

110. **Funder**: Very good where are you from
111. **Darya**: Me Ukraine
112. **Funder**: You're from the Ukraine how long have you been here in Dun Cluain
113. **Darya**: In Dun Cluain from July
114. **Funder**: From July living here that long and in Ireland
115. **Darya**: In February two years
116. **Funder**: Oh you are in Ireland two years okay . you learned English here in Ireland

It is the fluency of the interaction between lines 94–102 when she supports him and the fluency of lines 103–109 when he supports her, culminating in a quasi-closing by her (okay↑ on line 109 with rising intonation and prominence of the last syllable) which gives this 'impression'. This is corroborated by field notes of the encounter with Darya on this occasion, which reflected that I thought her level of English was very good and I would have, as an organizer of classes, placed her in a higher level class. However, looking back over her discourse (or more appropriately her 'language'—in the context of EFL assessment), there are many language mistakes (in the Chomskyan sense of 'well-formed sentences') that would impede Darya from obtaining a high score even on the most user friendly of assessment tools such as the Oxford Quick Placement test. Even in these 'fluent' sections she says:

(103) 'tomorrow I need going to town'
(105) 'and I talk with this girl'
(107) 'and she tell me price…'
(109) 'and I am back…'

The second example comes from the discourse of Dragana. She is requesting funding for her two daughters, but by the end of the 20-min meeting she has managed to get tacit agreement from the funder to pay for sporting activities for the two daughters and her husband, in addition to funding for the taxi that takes her daughter to basketball training on Saturday mornings. Dragana exploits *yeah* to the maximum: she uses this response token 57 times, which equates to 12.98% of her total discourse, which is notably greater than any other resident. However, 'okay', 'yes', 'you know' and 'no' used as discourse markers and the pause filler 'er' are all prominent as well, and all together these make up 28.7% of her discourse. These and her varied intonation contours contribute greatly to her negotiating her way through the interaction and creating an impression of fluency. The extract comes from the end of the meeting, beginning with her announcing, with relief palpable in her voice, that she is going for her tea.

Extract 6

149. **Dragana**: I now go for my tea
150. **Funder**: Okay no for Monday
151. **Dragana**: Okay okay <even intonation here>
152. **Funder**: And on Monday at four
153. **Dragana**: Okay ∨
154. **Funder**: Can you get her to no Monday no Wednesday
155. **Dragana**: Okay ↓ I know
156. **Funder**: It's Next Wednesday at four o'clock
157.➔ **Dragana**: Yeah ↓
158. **Funder** You get her to come here to speak to me
159. **Dragana**: Yeah↑ <surprise> <high pitch> okay∨ <high pitch>
160. **Funder**: or Monday, Monday at four o'clock
161.➔ **Dragana**: Okay↑ <High pitch> yeah ∨ <High pitch>
162. **Funder**: Monday at four
163.➔ **Dragana**: Yeah↑<high pitch> okay∨ <high pitch>
164. **Funder**: I'll tell you then
165. **Dragana**: Okay <high pitch>∨
166. <P10 while **Funder** writing>
167. **Funder**: Okay
168. **Dragana**: Yeah↑ <high pitch>
169. **Funder:** Monday at 4 pm
170. **Dragana**: Yeah↓ I know
171. **Funder**: And you the receipt
172.➔ **Dragana**: Yeah↓
173. **Funder**: You clip it on here
174. **Dragana**: Okay <even>
175. **Funder**: Not this now
176.➔ **Dragana**: yeah ↑ <high pitch>
177. **Funder**: the receipt for the forty euros
178.➔ **Dragana**: Yeah ∨ yes of course
179. **Funder**: Okay
180.➔ **Dragana**: Yeah↓

In this extract, a minimal response constitutes each of the 15 responses uttered by Dragana, as an appointment for one of the daughters and the funder with regard to the receipt for €40 is arranged. The responses are isolated in Table 17.

The complete gamut of meanings or functions with which Dragana infuses these response tokens are isolated in Table 18. This stretch of talk is representative of her whole discourse during the meeting and her exploitation of response tokens. While, the interpretation of attitude, functions, emotion or meaning of prosody is for the most part subjective, and I have largely interpreted it from an anglocentric point of view, the very infusion of the response tokens with the mixture of different contours, is evidence that there must be different values attached, whatever they may be, and that such diversity constitutes a subconscious concern to camouflage linguistic deficiency, reinforce the impression of fluency and negotiate meaning and objective. This concern is especially noticeable between lines 159 and 165 when Dragana and the funder are arranging the meeting for the daughter and the former uses a combination of prosodic features (intonation, loudness and pitch) to vary the

Table 17 Isolated utterances

Okay okay
Okay
Okay I know
Yeah
Yeah
Okay yeah
Yeah okay
Okay
Yeah
Yeah I know
Yeah
Okay
yeah
Yeah yes of course
Yeah

Table 18 Isolated response tokens, function and prosody

Line	Response tokens	Function	Prosody
151	*Okay okay*	Confirm/signal to continue	Even
153	*Okay*	Signal to continue	∨
155	*Okay I know*	Signal to continue, DM	↓
157	*Yeah*	Signal to continue	↓
159	*Yeah*	Surprise, acknowledge	↑ Loud High Pitch
159	*okay*	Signal to continue	∨ High pitch
161	*Okay*	Confirm	↑ High pitch
161	*yeah*	Signal to continue	∨ High pitch
163	*Yeah*	Acknowledge	↑ High pitch
163	*okay*	Confirm	∨ High pitch
165	*Okay*	Acknowledge	∨ High pitch
168	*Yeah*	Express understanding and close	↑ High pitch
170	*Yeah I know*	Signal to continue	↓
172	*Yeah*	Signal to continue/hasten	↓
174	*Okay*	Signal to continue	Even
176	*yeah*	Express surprise or acknowledge new information	↑
178	*Yeah*	Signal to continue	∨
178	*Yes*	Confirmation	Even
180	*Yeah*	Signal to continue and close	↓

use of the response tokens. Noticeable also is the use of the complex (see Tottie 1991) 'Yeah I know' (line 170) and 'Okay I know' (line 155) with the same intonation pattern on the response tokens, and the same intonation falling on the first

and second words of the complex response tokens 'yeah↑ okay∨' (159 and 'okay↑ yeah∨' (161), even when the first and second words are interchanged.

With regard to the prosodic features (Table 18) it could be said, in general that falling intonation is used for continuers, rising intonation is used for acknowledgement and rising intonation and loudness is used for surprise and acknowledgement of new information. High pitch seems to be reserved for utterances which express surprise.

The next example comes from the meeting with the two Ukrainians, Katya and Anyeta. In Extract 7. below, both 'yeah' (lines 6 and 13) are answers to polar questions, the first asked with the auxiliary 'did' and the second with 'would'. The second 'yeah' also prefaces a topic-switching turn, (see Jefferson 1983, 1993 for this use of *yeah*), Katya requesting funding to go with her baby, using the hedged performative'wanted' and hedges 'okay' 'maybe' and 'like' for further mitigation.

This is a good example of how one of the residents exploits her pragmatic knowledge, the multi-functional 'yeah', the hedged performative and other hedges, to arrive at where she was heading all along in the conversation—to establish if she can take her baby (who is not mentioned in this utterance) swimming. She has managed to negotiate the conversation in such a way that the funder's final comment in this fragment

'Take a form with you and fill it in for me and find …' (line 14)

conveys to her that she *can* take her baby.

The use of 'wanted' to mitigate the request in light of her non-marking of the past in other utterances also demonstrates the pragmatic priority of her language. The use of unmarked 'I ask' for *I asked* in line 2 below (for example) does not interfere with the referential and past meaning as it is embedded in the context, so it does not seem to be a source of concern for the speaker.

Extract 7

1.		**Funder:** Somebody asked me for swimming
2.	➔	**Katya:** I ask you but I'm busy I'm not I want not for myself I want for my son but small
3.		Michael to go with him but it's not kind of sports you see<unintelligible> she last year was
4.		with him a few times Jury's hotel
5.		**Funder:** Did you go?
6.	➔	**Katya:** Yeah (3)
7.		**Funder:** and
8.		**Katya:** It was okay they said like one hundred sixty for four months and five months you get
9.		free
10.		**Funder:** How much
11.		**Katya:** Hundred sixty euro
12.		**Funder:** And would you be interested in it again?
13.	➔	**Katya:** Yeah (4) but I wanted to ask you is it okay or is it like kind of maybe not
14.		**Funder:** Take a form with you and fill it in for me and
15.		**Aneta**: {ya
16.		**Funder**: and find

The word frequency lists show how dependent the language is on response tokens, especially on *yeah* and *okay* (which account for 10% of their output) in the Funding Corpus. The isolation of long stretches of turns here in which no other word is uttered, is further evidence of this dependency. The detailed analysis of how the residents use these response tokens, supporting Fellegy's (1995: 196) claim for them as the 'hinge between the linguistic and communicative competence,' completes the picture of their interactional success. In Table 18 above, various functions of *yeah*

accompanied by diverse prosodic patterns were listed, but these are only a portion of the functions identified in the larger study which include the use of:

- *Um*, *yeah*, *uh huh*, *oh* and *okay* to avoid the turn.
- *Yeah* to interrupt and to get funder to change topic.
- *Yeah* and *okay* as a discourse marker of speaker incipiency.
- *Okay* and *yeah* to introduce and change topic in much the same way as *right.*
- *Yeah* during background conversation in a different language to maintain lingua franca contact with a foreground speaker.
- *Ah* during conversation in a foreign language as a receipt of new information, but directed at the native speaker of English.
- *Yeah* as a hedge and as a strategy of politeness.
- *Yeah* as a continuer of extended turns.
- *Yeah* to pass the turn back quickly.
- *Yeah* as a hastener.
- *Yeah* for agreement.
- *Yeah* for tag questions in substitution of the canonical format.
- *Okay* and *yeah* to substitute semantically for both single words (polar *yes*) and long phrases.
- *Yeah* to express surprise.
- *Yeah* to substitute small talk.
- *Yeah* and *okay* to transmit boredom, disinterest and impatience.
- *Yes* for confirmation, and reinforcement of *yeah* and to preface a full turn.
- *Yeah* and *okay* are used as part of a sidestep topic shift process

Apart from these discrete uses of *yeah,* this item was found to change procedurally from minimal response to discourse marker according to the 'status' of the resident in the ongoing interaction. The default 'status' of the resident is passivity, not only in the transactional sense of the funder being the holder of the purse strings in this particular instance, but also in the deeper sense of power relations, the interlocutor (the funder) representing the powerful authority on which the asylum seeker is dependent for daily survival and future citizenship. Although the residents rarely initiated and chiefly used *yeah* as a continuer when the funder was explaining rules and regulations or was engaged in a long telling, they also used it as a discourse marker in place of more typical items such as *right* and *well* (Sinclair and Coulthard 1975; Carter and McCarthy 2006; O'Keeffe et al. 2007) when they took the floor, perhaps unconsciously dissimulating the shift in the power relationship. In such cases y*eah* was uttered with a different intonational contour and intensity.

The cumulative effect of the use of the minimal responses and *yeah* in particular served to camouflage disfluency and linguistic limitations, and perhaps camouflage underlying power over the trajectory of the talk by the residents. On the one hand, the *yeah*s signal to the native-speaking funder that he is in control and that the non-native speaker is at least understanding; the real effect is that the residents, almost surreptitiously, switch topics to personal requirements, which is the principal interest of the asylum seekers: the objective is successful negotiation with the English-speaking staff in order to acquire means for the satisfaction of basic needs.

The pragmatics of the interaction is more important for the residents than referential talk as is shown by the example of the non-use of the past simple form for time reference use in contrast its manipulation for mitigation.

Conclusions and Implications

In this paper I have shown how the corpus analysis of the lingua franca discourse of a closed community of speakers of other languages revealed that the members chiefly interact with 100 lexical and non-lexical words and that there is limited use of the components of what McCarthy (1999: 240) classified as the 'nine broad categories of a basic spoken vocabulary'. The question of how they interact with any success with the English-speaking staff of the centre, especially when transacting or negotiating basic needs and financial support, with such a limited vocabulary, arises. The word frequency lists of the lingua franca corpus revealed a high incidence of the minimal response *yeah,* and this pattern and the predominance of *yeah* and other minimal responses was even more striking in a smaller corpus of funding application meetings in which the residents of the centre negotiate financial support for participation in sporting activities. Closer analysis of these items in the here and now process of interaction revealed that the minimal responses, *yeah,* in particular, are exploited *ad maximum* semantically and discoursally by the residents, and that they are also used strategically to camouflage, consciously or unconsciously, any disfluency and linguistic deficiency that might interfere with the pragmatic enterprise of getting something done, or as is in this specific context, of getting approval for personal funding.

The implications of these findings are diverse but significant. Firstly, as far as methodology is concerned the blend of the ethnography, quantitative corpus linguistics and the qualitative detailed analysis of conversation has illuminated the lingua franca talk. I have not elaborated on the ethnography here, but in the original study it provided the initial contextual clues with regard to communicative patterns. The corpus analysis then, as shown here, provided an illuminating quantitative characterization of the lingua franca, but also raised an important question with regard to the 'difficulty' of communicating with such a vastly reduced language system. The interplay between quantitative and qualitative analysis facilitated the elucidation of this issue.

Secondly, the hard-scientific fact of the reduced core vocabulary, the absence of common components of spoken English, the preponderance of minimal responses (especially *yeah*) in the main corpus and funding corpus in particular, and the success in interaction, as shown by the qualitative analysis, despite such a priori limitations, casts doubt on the laboured acquisition of large lists of vocabulary by learners and the misguided focus of some teachers on 'teaching' mass vocabulary. This is especially relevant in the field of the provision of English to Speakers of Other Languages (ESOL), the learners usually being recently arrived immigrants striving to survive. In my own context as the manager of public ESOL provision and the director of a certificate in English language training (CELT), for example, trainee tutors ask how many words they should teach per day and how they should

teach them, and experienced ESOL teachers are accustomed to handing out long lists of vocabulary that the students will never use—in their quest for survival at any rate. As Sinclair and Renouf (1988 142–143) say, 'there is far more general utility in the recombination of known elements than in the addition of less easily usable items.' The simple lesson for teachers and students is that vocabulary acquisition and learning depends on community and context.

Thirdly, public service ESOL policies (in the context of Ireland here) need to consider such research when deciding whether to provide basic A1 and A2 instruction, as they generally do at the moment, or provide courses all the way to 'proficiency'. As this study shows that people can get by using a limited vocabulary, perhaps the provision should focus on these basic levels with the inclusion of firstly, a pragmatic element on the syllabus, and secondly input on learner autonomy in order that the immigrant students themselves further develop their English for more sophisticated use.

Fourthly, as has been seen here, the study of a lingua franca for its own sake brings into relief what Firth (1996) calls 'contingent' interactional achievement (see Harrington 2016, 2018). The benefit of the fact that English is used as the base language for massive lingua franca communication in the modern world, rather than feeding futile and at times antagonistic debate with regard to the entity of English as a Lingua Franca as a language variety in its own right,[3] should be seen in the light of the opportunity such mass but varied usage provides for researchers interested in the real mysteries of language, the real remit of linguists.

Finally, while most theories of communicative competence emerged as a response to the so-called limiting aspect of Chomsky's research remit (1965), they are themselves [see, for example, Hymes (1972), Canale and Swain (1980), Canale (1983), Bachman (1990) and Savignon 1997] restrictive in that they are constructs of 'competence' in a certain language or languages, in a certain culture or cultures, and in a certain community or communities, because they all prioritize the 'sharedness' of components. The analysis of the lingua franca here, turns the focus toward how people who do not have access to a shared language or culture, co-construct meaning in talk-in-interaction in an instinctive way. The lacuna in existing theories of communicative competence then is the consideration of non-language or non-culture-centred instinctive interactional competence—the procedural management of interaction in a foreign (in the sense of other) language, facilitated by whatever resources are on hand—minimal responses and prosodic features and the lingua franca instinct.

Compliance with Ethical Standards

Conflict of interest The author states that there is no conflict of interest.

[3] Proponents of this view include Seidlhofer (2010), and Jenkins (2009); while Sobkowiak (2005), Rubdy and Saraceni (2006), and Prodromou (2006) constitute the main opposition.

Appendix: Transcription symbols

↑ Rising intonational contour
↓ Falling intonational contour
∧ Rise-fall
∨ Fall-rise
◦ Soft intensity
Capitalization = Loud intensity
Underline = Prominence
<Syll > = one syllable uttered (and continuation of the word syllable to show duration)
<Sylla> = two syllables uttered
<Syllables> three syllables uttered
<Four syllables> four syllables uttered
<Unintelligible string> = more than four syllables uttered
{ = interruption
. = micropause
... = pause of circa one second
<2S> = two second pause, and so on.
hh = breath intake
haha = laughter (and repeated syllables to show the duration)
<happy> = the transcriber guesses the word uttered
∩ = latching
= utterances connecting meaning over various lines

References

Bachman, B. (1990). *Fundamental considerations in language testing*. Oxford: Oxford University Press.
Brazil, D. (1995). *A grammar of speech*. Oxford: Oxford University Press.
Bublitz, W. (1988). *Supportive fellow-speakers and co-operative conversations*. Amsterdam: John Benjamins.
Canale, M. (1983). From communicative competence to communicative language pedagogy. In J. Richards & R. Schmidt (Eds.), *Language and communication* (pp. 2–27). London: Longman.
Canale, M., & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics, 1,* 1–47.
Carter, R. (1998). *Vocabulary: Applied linguistics perspectives*. London: Routledge.
Carter, R., & McCarthy, M. J. (2006). *Cambridge grammar of English: A comprehensive guide*. Cambridge: Cambridge University Press.
Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.
Cogo, A., & Dewey, M. (2012). *Analysing English as a lingua franca: A corpus-driven investigation*. London: Continuum.
Dittman, A. T., & Llewellyn, L. G. (1968). Relationship between vocalisations and head nods as listener responses. *Journal of Personality and Social Psychology, 9,* 79–84.
Fellegy, A. M. (1995). Patterns and functions of minimal response. *American Speech, 70*(2), 186–199.

Firth, A. (1996). The discursive accomplishment of normality. On lingua franca English and conversation analysis. *Journal of Pragmatics, 26,* 237–259.
Fries, C. C. (1952). *The structure of English*. New York: Harcourt, Brace.
Gardner, R. (1998). Between speaking and listening, the vocalisation of understandings. *Applied Linguistics, 19*(2), 204–224.
Gardner, R. (2001). *When listeners talk: Response tokens and listener stance*. Amsterdam: John Benjamins.
Harrington, K. (2016). *Anything but the ideal speaker listener in anything but the homogeneous speech community*. Unpublished thesis, Mary Immaculate College.
Harrington, K. (2018). *The role of corpus linguistics in the ethnography of a closed community*: *Survival communication*. London: Routledge.
Heritage, J. (1984). A change-of-state token and aspects of its sequential placement. In J. Maxwell Atkinson & J. Heritage (Eds.), *Structures of social action: Studies in conversation analysis* (pp. 299–345). Cambridge: Cambridge University Press.
Hymes, D. (1972). On communicative competence. In J. Pride & J. Holmes (Eds.), *Sociolinguistics* (pp. 269–293). Harmondsworth: Penguin.
Jefferson, G. (1983). Notes of a systematic development of the acknowledgement token 'yeah' and 'mmhm'. *Papers in Linguistics, 17*(2), 197–216.
Jefferson, G. (1993). Caveat speaker: Preliminary notes on recipient topic-shift implicature. *Research on Language and Social Interaction, 26,* 1–30.
Jenkins, J. (2000). *The phonology of English as an international language*. Oxford: Oxford University Press.
Jenkins, J. (2009). English as a Lingua Franca: Interpretations and attitudes. *World Englishes, 28*(2), 200–207.
Kendon, A. (1967). Some functions of gaze-direction in social interaction. *Acta Psychologica, 20,* 22–63.
Labov, W., & Fanshel, D. (1977). *Therapeutic discourse: Psychotherapy as conversation*. New York, NY: Academic Press.
Lyons, J. (1971). *Introduction to theoretical linguistics*. Cambridge: Cambridge University Press.
McCarthy, M. J. (1999). What constitutes a basic vocabulary for spoken interaction. *Studies in English Language and Literature, 1,* 233–249.
Nation, P. (2006). How large a vocabulary is needed for reading and listening? *Canadian Modern Language Review, 63,* 59–82.
O'Keeffe, A., & Adolphs, S. (2008). Using a corpus to look at variational pragmatics: Response tokens in British and Irish discourse. In K. P. Schneider & A. Barron (Eds.), *Variational pragmatics* (pp. 69–98). Amsterdam: John Benjamins.
O'Keeffe, A., McCarthy, M. J., & Carter, R. A. (2007). *From corpus to classroom: Language use and language teaching*. Cambridge: Cambridge University Press.
Prodromou, L. (2006). A reader responds to J. Jenkins's "Current perspectives on teaching World Englishes and English as a Lingua Franca". *TESOL Quarterly, 41*(2), 409–413.
Prodromou, L. (2008). *English as a lingua franca*. London: Continuum.
Rubdy, R., & Saraceni, M. (Eds.). (2006). *English in the world: Global roles*. London: Continuum.
Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A simplest systematics for the organisation of turn-taking for conversation. *Language, 50*(4), 696–735.
Savignon, S. (1997). *Communicative competence: Theory and classroom practice*. New York: McGraw-Hill.
Schegloff, E. A. (1972). Sequencing in conversational openings. In J. Fishman (Ed.), *Advances in the sociology of language, II: Selected studies and applications* (pp. 91–125). The Hague: Mouton.
Schegloff, E. (1982). Discourse as interactional achievement: some uses of 'uh huh' and other things that come between sentences. In D. Tannen (Ed.), *Analysing discourse: Text and talk* (pp. 71–93). Washington: Georgetown University Press.
Schiffrin, D. (1987). *Discourse markers*. Cambridge: Cambridge University Press.
Schmitt, N., Cobb, T., Horst, M., & Schmitt, D. (2015). How much vocabulary is needed to use English? Replication of van Zeeland & Schmitt (2012), Nation (2006) and Cobb (2007). *Language Teaching*. doi:10.1017/S0261444815000075.
Seidlhofer, B. (2010). Giving VOICE to English as a Lingua Franca. In R. Facchinetti et al. (Eds.), *From international to local English—and back again* (pp. 147–163). Frankfurt: Peter Lang.
Sinclair, J. M., & Coulthard, M. (1975). *Towards an analysis of discourse*. Oxford: Oxford University Press.

Sinclair, J. M., & Renouf, A. (1988). A lexical syllabus for language learning. In R. Carter & M. McCarthy (Eds.), *Vocabulary and language teaching* (pp. 140–158). London: Longman.

Sobkowiak, W. (2005). Why not LFC? In K. Dziubalska-Kołaczyk & J. Przedlecka (Eds.), *English pronunciation models: A changing scene* (pp. 131–151). Frankfurt am Main: Lang.

Stubbe, M. (1998). Are you listening? Cultural influenceson the use of supportive verbal feedback in conversation. *Journal of Pragmatics*, *29*, 257–289.

Tottie, G. (1991). Conversational style in British and American English, The case of backchannels. In K. Aijmer & B. Altenberg (Eds.), *English Corpus Linguistics* (pp. 254–273). London: Longman.

van Zeeland, H., & Schmitt, N. (2012). Lexical coverage in L1 and L2 listening comprehension: the same or different from reading comprehension? *Applied Linguistics*. doi:10.1093/applin/ams074.

Yngve, V. (1970). On getting a word in edgewise. *Papers from the 6th regional meeting, Chicago Linguistic Society*. Chicago: Chicago Linguistic Society (pp. 567–577).