“*I don’t know*”: a usage-based approach to familiar collocations in non-fluent aphasia

Claudia Bruns[1], Rosemary Varley[1], Vitor C. Zimmerer[1], Marcella Carragher[2], Gwen Brekelmans[1], & Suzanne Beeke[1]

[1] *Department of Language & Cognition, Division of Psychology and Language Sciences, University College London, UK*
[2] *Discipline of Speech Pathology, Department of Community & Clinical Allied Health, La Trobe University Melbourne, Australia*

Corresponding author: Claudia Bruns
Department of Language & Cognition, Division of Psychology and Language Sciences, University College London, Chandler House, 2 Wakefield Street, London WC1N 1PF
E-mail: c.bruns@ucl.ac.uk

Abstract

Background: Familiar collocations (e.g., "*it's alright*") are an important part of everyday conversation. Such word combinations are often retained in speakers with Broca's aphasia. However, only few investigations have studied the forms and functions of familiar collocations available to speakers with Broca's aphasia.

Aims: We first apply a frequency-based perspective to word combinations produced by speakers with Broca's aphasia and their conversation partners (CPs), and compare the frequency characteristics of word combinations in dyadic and non-dyadic speech. Second, we investigate the conversational functions of one prominent familiar collocation, "*I don't know*" (IDK).

Methods & Procedures: In the first analysis, speech samples from interactions of nine dyads (each a speaker with Broca's aphasia and their CP) were examined. Non-dyadic samples were selected from 39 speakers with Broca's aphasia from AphasiaBank (MacWhinney et al., 2011). The *Frequency in Language Analysis Tool* (FLAT; Zimmerer & Wibrow, 2015) was used to estimate collocation strength (the degree of association between words in a combination) of well-formed bigrams (two-word combinations) and trigrams (three-word combinations). The second analysis presents a qualitative investigation of uses of IDK in dyadic exchanges.

Outcomes & Results: Analysis 1 revealed that residual trigrams in Broca's aphasia were more strongly collocated in comparison to language produced by CPs. There was no difference in frequency-based profiles between dyadic and non-dyadic aphasic speech. Analysis 2 indicated that speakers with Broca's aphasia and CPs used IDK to achieve a variety of communicative functions. However, patterns specific to each participant group were found.

Conclusions: These findings highlight that frequency-based analysis is useful in explaining residual, grammatically well-formed word combinations in Broca's aphasia. This study provides evidence that IDK can aid turn construction in aphasia.

Keywords:

Broca's aphasia, frequency-based, collocations, formulaic expressions

Introduction

Despite grammatically impoverished, non-fluent speech in Broca's aphasia, residual multi-word utterances (e.g., "*I don't know*") are often fluently produced and employed in appropriate situations. Such multi-word utterances often represent common word combinations, some of which are referred to as formulaic expressions. The latter term describes a prefabricated sequence of words, as well as single words such as the discourse particle "*well*" (Van Lancker Sidtis & Postman, 2006; Wray, 2002). They make up a large proportion of normative conversation and are familiar to native speakers. Estimates vary from 24-48% (Van Lancker Sidtis & Rallon, 2004; in unscripted telephone conversations) to 59% of typical conversation (Erman & Warren, 2000). Formulaic expressions represent a significant part of the conversational inventory, and are a feature of proficient language use. They can be grouped into idiomatic (e.g., "*it's raining cats and dogs"*) and non-idiomatic phrases (e.g., "*I know what you mean*") and may be fully or partly fixed in form (e.g., "*and so on*"; "*could I have [X]*"). In neurotypical talk, formulaic expressions are associated with various conversational functions. For example, the expression "*I see*" is used as a backchannel or feedback signal (Erman & Warren, 2000; Schmitt & Carter, 2004). In this study, we use corpus methods to establish the collocation strength, or degree of association between components of multi-word expressions used by speakers with aphasia and their conversation partners (CPs). Word combinations which are more frequent and more strongly collocated, are more likely formulaic.

While the forms and functions of familiar collocations in neurotypical talk are well studied (e.g., Conrad & Biber, 2004; Drew & Holt, 1998; Kecskes, 2000; Meunier, 2012; Schmitt & Carter, 2004), systematic investigations of their use in aphasia are relatively rare. Code (1982) analysed recurrent utterances (RUs) across subtypes of aphasia. Code found real-word RUs mostly consisted of high-frequency words (e.g., "*I told you*", "*so so*") and typically occurred in Broca's aphasia. A salient pattern was the pronoun + verb group (e.g., "*I want to*"). These forms are often described as stereotypes or lexical automatisms (Blanken, 1991; Code, 1982; Grande et al., 2008), suggesting pathological behaviour (Blanken & Marini, 1997; Rodrigues & Castro-Caldas, 2014). However, as typical language consists of large proportions of familiar collocations, whether they should be viewed as markers of pathology in Broca's aphasia is unclear.

Van Lancker Sidtis & Postman (2006) explored the neural basis of formulaic expressions. They recorded the use of numerals, proper nouns, idioms, conventional expressions (e.g., "*as a matter of fact*"), speech formulas (e.g., "*right*"), expletives (e.g., "*damn*"), sentence stems (e.g., "*I guess*"), discourse particles (e.g., "*well*") and pause fillers (e.g., "*uh*") in spoken output of neurotypical individuals, speakers with fluent aphasia, and non-aphasic individuals with right hemisphere damage (RHD). They found higher proportions of formulaic expressions in speakers with aphasia compared to individuals with RHD and suggested that these expressions are represented in the right hemisphere (Sidtis, Canterucci, & Katsnelson, 2009). Other investigators have explored idiom comprehension across participants with aphasia and RHD (Van Lancker & Kempler, 1987), and the processing of "automatic language" (counting from 1 to 10; completing familiar, idiomatic phrases; repeating idiomatic phrases) in fluent and non-fluent aphasia (Lum & Ellis, 1999). The findings indicate relatively preserved production and comprehension of these expressions, many of them idiomatic, in aphasia.

Familiar collocations, due to their high usage frequency, can be explained by theories such as usage-based Construction Grammar (Goldberg, 2003). This approach differs from traditional generative grammatical theories in that linguistic knowledge is assumed to be represented as constructions, where form is paired with meaning. Constructions vary in size and complexity, i.e., they can be multi-word utterances. A number of investigators have applied this approach to phenomena in aphasic language (e.g., DeDe, 2013; Gahl & Menn, 2016). Zimmerer, Newman, Thomson, Coleman, & Varley (2018), for example, analysed language production in semi-structured interviews by speakers with fluent aphasia, non-fluent aphasia and RHD from a frequency-based perspective. They showed that speakers with aphasia, but not speakers with RHD, relied on more strongly collocated word combinations.

While familiar collocations are known to be retained in aphasia, it remains unclear whether such combinations are a main feature of aphasic language production in everyday conversations, and whether they perform specific conversational functions in aphasia. The first analysis of the current report adopts a frequency-based perspective to word combinations. We establish frequency characteristics at the level of two-word- (bigrams) and three-word combinations (trigrams). We employ automated analysis software, the *Frequency in Language Analysis Tool* (FLAT; Zimmerer & Wibrow, 2015; Zimmerer, Wibrow, & Varley, 2016; Zimmerer et al., 2018), to analyse language produced by speakers with aphasia in different talk contexts. The FLAT determines usage frequency of every word and word combination in a test sample using the 10 million word spoken-conversation section of the British National Corpus (BNC, 2007), which represents typical language use. In addition to frequency of occurrence, the FLAT automatically computes collocation variables of bi- and trigrams. We employ a commonly used measure of collocation strength, the t-score. For instance, the trigram "*it's* *alright*" has a t-score of 28 (based on the spoken BNC), whereas "*it's new*" has a t-score of 4. Higher values point to more strongly associated or collocated word combinations, reflecting a higher certainty of collocation. In this way, in analysis one, we explore the following research questions: a) How strongly associated are the word combinations produced by speakers with aphasia compared to CPs in naturalistic conversations? b) Do frequency-based values vary across talk contexts in aphasia? We hypothesize that speakers with aphasia rely on more common multi-word utterances as compared to neurotypical control speakers, and that the frequency-based values are robust across different talk contexts.

Analysis 2 focuses on familiar collocations in interaction. Conversation analytic studies show that such expressions can be interactionally beneficial in aphasia. For example, the "*I suppose*" construction (Beeke, 2003) was produced by a speaker with chronic Broca's aphasia five times in a 13 minute conversation to express his opinion. Familiar collocations have also been viewed as compensatory strategies in aphasic talk. Simmons-Mackie & Damico (1997) found that set phrases like "*very nice*" and "*all the time*" – although they might not add new information to the conversation – were used as a resource by two speakers with Broca's aphasia to regulate interaction, or in the case of "*all the time*", to express magnitude. The second analysis explores the conversational uses of "*I don't know*", which, in our dataset, was common in non-fluent aphasic talk and control participants.

Method

Participants

This study used pre-collected everyday conversation data (dataset 1) and semi-structured interview data (dataset 2). Conversation data stem from nine dyads, each comprising a person with post-stroke non-fluent aphasia (at least 6 months post-onset) and their neurotypical CP.

Dyads 1-8 took part in a study by Carragher, Sage, & Conroy (2013), and dyad 9 in a study by Best et al. (2016)[1]. Background information on all dyads is summarized in Table 1. All participants presented with Broca's aphasia. Aphasia classification was based on clinical consensus, grammatically impoverished output on picture description tasks, and performance on standardised language assessments (Table 1; for more details see Carragher et al., 2013, p. 852; Beeke, Maxim, Best, & Cooper, 2011, p. 228). Dyads took part in intervention studies (dyads 1-8: Carragher et al., 2013; Carragher, Sage, & Conroy, 2015; dyad 9: Best et al., 2016) and recorded weekly conversations prior to, during and after intervention. Analysis is based on pre-therapy recordings only and the nature of interventions is not relevant to the current study. Participants gave written informed consent to long term storage of their data. Ethical approval was granted by NHS IRAS ethics (Carragher et al., 2013) and NHS Cambridgeshire 1 Research Ethics Committee (Best et al., 2016).

(Table 1 near here)

[1] Dyad 2 in Best et al.

Table 1. Naturalistic interaction participant information.

Dyad	Initials	Corpus	Speaker with Broca's aphasia					CP relation to speaker with aphasia
			Gender	**Age at time of recording**	**Time post onset (months)**	**Naming objects (% correct)**	**Naming actions (% correct)**	
1	**KK**	Carragher et al., 2013	male	48	24	15	18	Wife
2	**GL**		male	47	12	32	22	Partner
3	**BL**		male	64	57	45	31	Wife
4	**DC**		male	40	72	32	30	Father
5	**JH**		female	36	8	27	43	Husband
6	**PM**		male	67	47	60	59	Wife
7	**PG**		male	66	132	65	65	Wife
8	**DM**		male	48	36	72	85	Wife
9	**SC**	Best et al., 2016	male	39	30	83	55	Wife
			Mean	**50.6**	**46.4**	**47.7**	**45.1**	
			SD	**12.1**	**38.1**	**23.2**	**22.3**	

Naming objects: Percentages are derived from the Boston Naming Test (BNT; Kaplan, Goodglass, & Weintraub, 2001), used for assessing lexical retrieval for dyads 1-8; for dyad 9, the percentage reflects an average across three baselines of the OANB (Druks & Masterson, 2000) noun subset.

Naming actions: Percentages are based on performance in the OANB verb subset (an average across two baselines for dyads 1-8 and three baselines for dyad 9).

Dataset 2 comprised semi-structured interviews from the online database AphasiaBank (MacWhinney, Fromm, Forbes, & Holland, 2011). AphasiaBank classifies aphasia subtype with the Western Aphasia Battery (WAB; Kertesz, 2007). Only participant 9 (SC) from dataset 1 had a WAB assessment and so AphasiaBank samples were selected that matched as closely as possible to that speaker's WAB profile (Broca's aphasia, Aphasia Quotient [AQ] of 60.7). Thirty-nine participants (13 female, 26 male) were selected from AphasiaBank (Table 2). They were classified as having Broca's aphasia and their mean age was 55.7 years (SD = 11.4). The AphasiaBank group had a WAB AQ between 50.0 and 70.4, indicating a range from moderate to mild aphasia. The mean time post-onset of the AphasiaBank group was 88 months (SD = 70; range: 6 to 309 months).

(Table 2 near here)

Table 2. Semi-structured interviews participant information (from AphasiaBank).

Participant #	ID	WAB AQ	Age *	Gender	Time post onset (months)
1	ACWT01a	63.9	70	F	142
2	ACWT05a	57.7	76	M	84
3	adler10a	51.2	45	M	23
4	adler13a	55.8	52	M	60
5	adler16a	57.2	64	M	82
6	BU07a	51.5	52	M	50
7	cmu02b	59.6	43	M	140
8	elman03a	66.2	55	M	132
9	elman11a	67	52	M	109
10	fridriksson12a	57.5	48	F	118
11	kansas09a	60	57	M	68
12	kempler03a	60.7	65	M	192
13	kempler04a	54.6	60	F	40
14	kurland13a **	54	55	M	19
15	kurland19a **	67.2	71	F	106
16	kurland24a **	51.5	47	M	11
17	kurland29a **	58.2	61	M	6
18	MSU05a **	68.2	73	M	86
19	MSU07a **	61.4	26	F	15
20	scale01a	52.5	78	M	309
21	scale10a	63.5	45	M	180
22	scale15b	57	59	M	58
23	scale18a **	60.9	50	F	183
24	scale26a	64.8	59	M	135
25	TAP11a	58.1	63	F	45
26	TAP14a	60.2	45	M	15
27	TAP17a	59.5	65	F	28
28	TAP19a	59.4	55	F	28
29	TCU03a	70.1	42	M	75
30	TCU08a	63.9	57	M	95
31	whiteside04a	50	65	M	272
32	whiteside08a	54.7	38	M	12
33	whiteside12a	54.3	70	F	70
34	williamson12a **	64.4	43	M	41
35	williamson19a	69.4	53	F	57
36	wright201a	57.6	55	M	36
37	wright205a	59.7	56	M	71
38	wright206a	53.7	39	F	143
39	wright207a	61.5	64	F	107
	Mean	***59.5***	***55.7***	-	***88.2***
	SD	***5.3***	***11.4***	-	***69.9***

*Note. * Age at first testing session; ** for these participants, more than one sample was available*

Datasets

With regard to dataset 1, there were forty pre-therapy videotaped conversation samples. These consisted of four samples recorded over 4 weeks for dyads 1-8, and eight samples over 8 weeks for dyad 9. All conversations were recorded in the dyads' home under instruction to videotape a conversation at a time of day and on a topic of the dyad's choosing. Conversations were transcribed as part of prior studies by Carragher et al. (2013) and Best et al. (2016), or by the first author of the current project where additional untranscribed samples existed. Sampling consisted of 5- to 25-minute segments where the participant with aphasia was in conversation with a family member. As a standard, the first 5-minute segment of a recording was not sampled, to allow participants to feel less conscious about the presence of the camcorder (for one dyad, some pre-transcribed samples started at the beginning of the recording). The videoed samples yielded 269 minutes of transcribed conversation which forms the basis for the current analysis. The average transcriptions per dyad reflected 22 min of conversation for dyads 1-8 and for dyad 9, 93 min of conversation.

Dataset 2 comprised 50 discourse samples from the 39 participants selected from AphasiaBank (one sample for 31 participants, and in order to maximise sample size, two to three samples from eight participants who were tested on multiple occasions). These reflect semi-structured interviews, where a clinician asks questions about a participant's speech, stroke story and an important event. The average duration of samples was 4:57 minutes (SD = 2:20). In total, 247 minutes of transcribed material from AphasiaBank were analysed.

Analysis 1

FLAT Version 1.1 was used for analysis. The program automatically extracts uni-, bi-, and trigrams in a sample of transcribed language, determines values on a number of measures of productivity, and derives frequency-related values from the spoken BNC for words and word combinations. For example, it segments the trigram “*it’s alright*” into three unigrams (“*it*”, “*‘s*”, “*alright*”) and two bigrams (“*it’s*” and “*‘s alright*”). Since only grammatical utterances are expected to appear in the spoken BNC, and ungrammatical utterances are atypical and would therefore represent very low frequency or non-existent combinations in a normative database, only grammatical strings are analysed by FLAT. Prior to frequency-based analysis, all transcripts were formatted in a manner consistent with the conventions of the FLAT (Zimmerer, Wibrow, & Varley, 2016; Zimmerer et al., 2018). For example, the utterance “*but seven days cycling*” was analysed as two separate clauses: “*but seven days*” and “*cycling*”. Clause boundaries were marked with separators (“<.>”). “<fill>” was used to replace any non-lexical interjections (e.g., “*erm*”), pauses and repetitions of words other than “*yes*”/ “*yeah*”, “*oh*” and “*no*”.

FLAT outputs also include measures of productivity: the number of word combinations, and combination ratio (the number of trigrams produced by a speaker divided by the number of words). Speakers with higher combination ratios display more output consisting of word combinations, i.e., better ability to combine words into multi-word expressions. More traditional indicators of lexical diversity are type-token ratio and vocabulary diversity (MacWhinney et al., 2011). However, these measures investigate productivity at the single word level, while the present study investigates the amount of connected speech produced. Hence, we report combination ratio.

To determine degree of association between words of the bi- and trigrams, we used t-scores, an association measure that indicates whether words co-occur more frequently than would be expected if all the words in the corpus were randomly distributed (Durrant & Doherty, 2010;

Gries, 2010; Hunston, 2002).[2] The trigram "*it's alright*" has a t-score of 28, whereas "*it's new*" has a t-score of 4. Hence, the words in "*it's alright*" are more strongly associated or collocated. Collocation strength is one marker of formulaic expressions (Zimmerer et al., 2018). We reported t-scores if a bi- or trigram had a frequency of occurrence of >1. Thus, our results are based on combinations that occur in the spoken BNC. FLAT outputs include both type and token bigram and trigram summary statistics. Type values reflect the inventory of bigrams and trigrams, while token measures reflect how frequently individual types are used. Table 3 shows details of FLAT variables and their calculation.

(Table 3 near here)

[2] The FLAT 1.1 employs additive smoothing by which 1 is added to every unigram, bigram and trigram frequency count in order to avoid a frequency value of 0 to enable calculation of t-scores.

Table 3. Frequency-related variables used in the FLAT analysis.

		Utterance type		
		Unigram (e.g., *it; 's; alright*)	**Bigram** (e.g., *it's*)	**Trigram** (e.g., *it's alright*)
Variables	**Observed absolute frequency**: occurrence in the spoken BNC	253864 (*it*) 199264 (*'s*) 7994 (*alright*)	68662	772
	Observed relative frequency: occurrence per million words	25386.4 19926.4 799.4	6866.2	77.2
	Expected frequency (taking into account the approximate number of words in the spoken BNC, 10,000,000)	-	$\frac{frequency_{it}frequency_{'s}}{10,000,000}$ (e.g., 5058.60)	$\frac{frequency_{it}frequency_{'s}frequency_{alright}}{10,000,000^2}$ (e.g., 4.04)
	t-score	-	$\frac{frequency_{it's} - expected\ frequency_{it's}}{\sqrt{frequency_{it's}}}$ (e.g., 242.7)	$\frac{frequency_{it's\ alright} - expected\ frequency_{it'salright}}{\sqrt{frequency_{it's\ alright}}}$ (e.g., 27.6)
	Combination ratio	Raw number of trigrams produced by a speaker divided by number of words		

Results

For group comparisons, Shapiro-Wilk tests were used to determine whether variables were normally distributed. Where variables were normally distributed, two-tailed independent t-tests were performed; otherwise, two-tailed Mann-Whitney U tests were employed. Pearson's r was calculated to report effect sizes.

Combination ratio was compared for the two speaker groups. The mean combination ratio for speakers with aphasia was .17 (SD = .06), and for CPs, it was .54 (SD = .03), representing a significant difference, t (16) = -15.16, p < .001, r = .97 (large effect). This confirms the agrammatic status of the speakers with aphasia, characterized by the limited ability to produce grammatical multi-word expressions.

How strongly associated are the word combinations produced by speakers with aphasia compared to CPs in naturalistic conversations (dataset 1)?

The t-score profiles for bi- and trigram types of both speaker groups are shown in Figure 1. For bigrams, the nine speakers with aphasia produced a total of 988 types (range: 30-262) and 1809 tokens (range: 101-413), and their CPs produced 7467 bigram types (range: 317-2082) and 11160 tokens (range: 414-3636). The average type-based bigram t-score for speakers with aphasia was 26 (SD = 7), and for CPs it was 19 (SD = 4), representing a significant difference between the two groups: t (16) = 2.71, p = .016, r = .56 (large effect). With regard to bigram tokens, the aphasic group had a higher average bigram t-score compared to the CP group (M = 53, SD = 23 versus M = 37, SD = 3, respectively). This difference was not significant, t (8.24) = 2.091, p = .069, r = .59 (large effect). A higher bigram t-score in the aphasic group, however, indicates that the constituents of the bigrams were more strongly associated with each other. The large effect size combined with a nonsignificant difference could be due to the variability within the aphasic group.

At the level of trigrams, speakers with aphasia produced 469 types (range: 11-172) and 698 tokens (range: 14-196). The CP group produced 5705 trigram types (range: 210-1751) and 6644 tokens (range: 236-2160). The effects observed at the bigram level were stronger for trigrams. For trigram types, the aphasic group displayed a higher average t-score of 17 (SD = 3) compared to 11 in the CP group (SD = 1). The difference was significant, t (9.864) = 5.74, p < .001, r = .88 (large effect). The mean token-based trigram t-score was 39 (SD = 29; median = 26) for the aphasic group, and 14 (SD = 1.9; median = 14) for the CP group. The difference was significant, U = 8.00, z = -2.87, p = .004, r = -.68 (large effect), indicating more strongly collocated word combinations in speakers with aphasia compared to neurotypical speakers. However, there was high inter-subject variability in the number of trigrams produced and average trigram t-scores in the aphasic group, with one speaker (GL) presenting with average token- and type-based trigram t-scores that were in the normative range (see Appendix 1).

(Figure 1 near here)

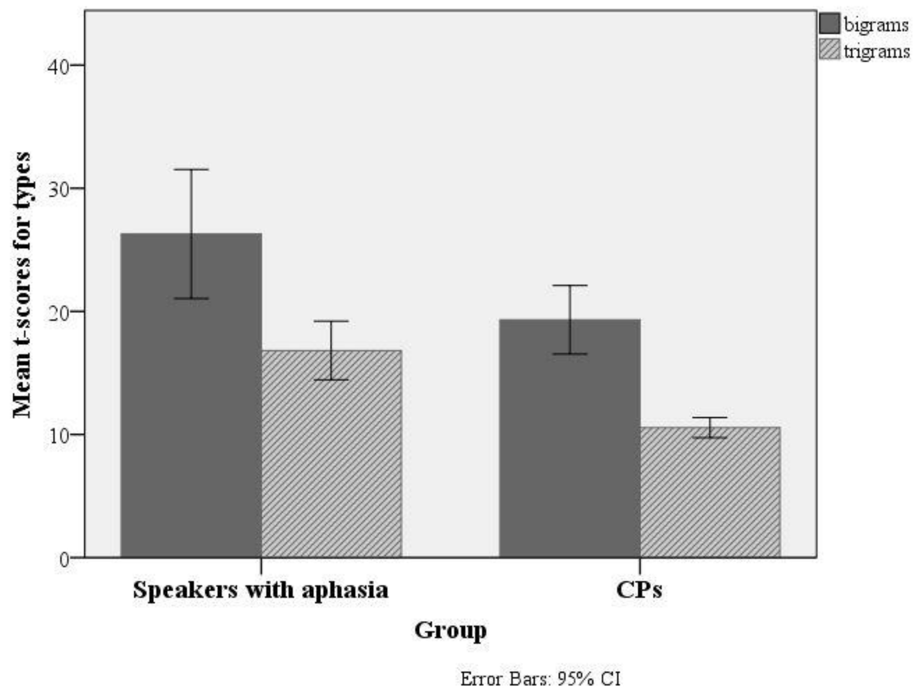

Figure 1. Average type-based bi- and trigram t-scores for speakers with aphasia and CPs.

The trigram types most frequently used by the nine speakers with aphasia are shown in Table 4. The "*don't know*" construction is a phrase available to seven out of nine speakers with aphasia (exceptions: DC, GL) and no other trigram was as widely used. In comparison, the three most frequently used trigrams in the CP group were: "*I don't*" (72 tokens, 1.08% of all CP tokens, used by all nine speakers), "*don't know*" (67 tokens, 1.01% of all CP tokens, used by all nine speakers), and "*do you want*" (26 tokens, 0.39% of all CP tokens, used by six speakers).

(Table 4 near here)

Table 4. Most frequently produced trigram types and tokens, everyday conversational data (speakers with aphasia).

Trigram type	**Total tokens produced**	**Proportion out of all trigrams**	**Trigram t-score**	**BL**	**DM**	**DC**	**GL**	**JH**	**KK**	**PM**	**PG**	**SC** (8 samples)
don't know	77	9.5%	95.01	21	3	0	0	21	17	1	1	13
I don't	72	8.8%	136.32	22	0	1	0	22	15	1	1	10
going to do	11	1.4%	26.94	0	0	0	0	0	0	11	0	0
one two three	10	1.2%	30.19	0	0	0	4	1	2	0	2	1
two three four	7	0.9%	26.44	0	0	0	3	1	1	0	2	0
it's alright	7	0.9%	27.64	0	0	0	0	0	0	0	0	7
that's it	6	0.7%	45.86	0	0	2	0	0	4	0	0	0
wait a minute	5	0.6%	16.67	0	0	0	2	0	3	0	0	0
n't know what	5	0.6%	39.33	0	0	0	0	0	5	0	0	0

Do frequency-based values vary across talk contexts in aphasia (dataset 1 vs. dataset 2)?

The average combination ratio of the AphasiaBank group was .25 (SD = .11), compared to .17 (SD = .06) in the conversational data, indicating that speakers in the AphasiaBank group were more successful in combining words into multi-word utterances. This group difference was significant, t (45) = -2.066, p = .045, r = .29 (small effect) which could stem either from the different conversational setting (elicited monologue in semi-structured interviews versus dyadic talk) or from varying aphasia severity across the two speaker groups.

AphasiaBank participants produced 1955 different bigram types and 4743 tokens. However, there was a high inter-individual variation: Participant #37 only produced 9 tokens, while participant #10 produced 535 tokens. Average type-based bigram t-scores in the AphasiaBank group (M = 31, SD = 13) did not differ significantly from the naturalistic data (M = 26, SD = 7), t (24.186) = -1.66, p = .109, *ns*, r = .32 (medium effect). The average token-based bigram t-scores in semi-structured interviews was 48 (SD = 26; compared to 53 (SD = 23) in conversational data). Again, a comparison of bigram t-scores revealed no significant difference, t (46) = .464, p = .645, *ns*, r = .07 (small effect). This indicates that the frequency-based characteristics with regard to collocation strength in the two settings are similar.

With regard to trigrams, the AphasiaBank group produced 1708 different trigram types and 2596 tokens. One participant (#37), however, did not produce any three-word combinations. A comparison of type-based trigram t-scores in the two speaker groups (AphasiaBank: M = 22, SD = 20, median = 16; naturalistic data: M = 17, SD = 3, median = 16) was not significant, U = 159.00, z = -.324, p = .746, *ns*, r = -.05 (small effect). The token-based trigram t-scores did not differ across datasets (everyday conversations: M = 39, SD = 29; semi-structured interviews: M = 30, SD = 27), U = 128, z = -1.16, p = .255, *ns*, r = -.17 (small effect). Again, the t-score measures appear robust across contexts, and t-scores from constrained elicitation conditions show good ecological validity.

Interim summary

The frequency-based analysis showed that speakers with aphasia, in comparison to their neurotypical CPs, employed significantly fewer word combinations. Furthermore, higher trigram t-scores (both type- and token based) in the aphasic group indicated that the words within combinations were more strongly associated. Use of more strongly collocated trigrams in speakers with aphasia appears stable across conversational settings. Combination ratio was lower in conversational samples than in semi-structured interviews. This might be due to more severe aphasic impairments in the dyadic group or the influence of probe questions designed to elicit extended monologue in the semi-structured interviews.

Despite high inter-subject variability with regard to specific constructions used in conversations, there was a small common subset across the nine speakers with aphasia including bigrams such as "*I know*", "*no no*", "*it's*", and the trigram "*don't know*". Moreover, the AphasiaBank samples included a further 89 "*don't know*" tokens. "*don't know*" was also produced at least once by each speaker in the CP group. The second analysis addresses conversational functions of these "*I don't know*" (IDK) expressions.

The functions of IDK

The word combinations produced by speakers with aphasia are more strongly collocated compared to those of their CPs. Analysis 2 investigates whether or not familiar collocations were used in a functionally typical way, based on one construction available to most speakers in the aphasic group: IDK. Typically, IDK or its reduced variant "*I dunno*", is associated with an inability to supply information (Hesson & Pichler, 2016). It often occurs in reply to a question, and is sometimes followed by a complement, as in "*I don't know* <u>*where he went*</u>". However, use of the phrase is not restricted to this prototypical meaning (e.g., Tsui, 1991; Diani, 2004; Pekarek Doehler, 2016). IDK can also function to avoid commenting, disagreement, or commitment to an answer. Other functions are marking uncertainty (hedging), and minimising compliments (Grant, 2010).

To our knowledge, there is currently no systematic investigation of the use and functions of IDK in aphasia. Hesson & Pichler (2016) analysed IDK use by speakers with dementia during cognitive assessment, and showed that severity of cognitive impairment was positively associated with the prototypical use of IDK. Mikesell (2009) explored it use in a single case study of an individual with frontotemporal dementia. She found that IDK often functioned as an appropriate answer to *wh*-questions, but sometimes occurred as a strategy to withdraw from a conversation and in other situations reflected memory difficulties. In the current analysis, the primary research question was: What are IDK usage patterns in speakers with aphasia when engaged in dyadic conversation, and do these patterns differ between aphasic and CP groups?

The analysis was based on dataset 1. We expected a higher proportion of isolated IDK tokens in aphasia, but did not have any other predictions regarding similarities and differences in the two speaker groups. Hence, our analyses are exploratory.

Analysis 2

IDK instances used by speakers with aphasia and CPs in dyads 1-9 were identified. Instances of "*don't know*" or "*dunno*", co-occurring with an explicit or implicit first-person pronoun "*I*", were included in the analysis. CP data were used as a normative sample of IDK usage. Video clips were extracted of each IDK example and one to two turns before and after the token to allow coding of conversational function.[3] IDK tokens were analysed separately for each speaker group.

All IDK tokens were coded for phonetic form, syntactic variation and conversational function using a rating system adapted from Pichler & Hesson (2016), Hildebrand-Edgar (2016) and Diani (2004). Four phonetic forms, the full "*don't know*" and the reduced "*dunno*", both with and without pronoun, were distinguished. For syntactic variations, there were five categories: isolated, IDK with *wh*-word, complement, co-occurrence with a discourse marker (e.g., "*well*", "*so*"), and "other" to capture constructions that could not be assigned to any of these categories. Conversational functions were coded using five main categories: lack of knowledge (LOK), interpersonal (INT), turn-constructional (TC), multifunctional (M; a combination of any of the mentioned functions) and "unclear function" (U) which was added for instances where there was not enough context or evidence to assign a category. Appendix 2 provides an overview and examples of the function rating system. For INT and TC, sub-categories were assigned to enable documentation of more specific functions such as avoiding commenting (sub-function of INT), or yielding the conversational floor (sub-function of TC; see Appendix 2). However, the five main categories were used to quantify the distribution of conversational functions.

Following the steps in Pichler & Hesson (2016), rater 1 (R1; first author) coded the video clips for conversational functions of all IDK tokens, while a second rater (R2; fifth author) independently coded all tokens for phonetic form and syntactic variation. Inter-rater reliability (IRR) was determined based on a random selection of a subset of 23 CP tokens (37% of data) and 29 tokens from speakers with aphasia (38% of data). These were coded by the other rater, i.e. R1 coded these for form and syntax, and R2 coded these for function.

[3] Unless the token occurred at the beginning or end of a sample or was followed by a long pause.

IRR was established on data from both groups. During the process of calculating IRR, tokens where one rater assigned a multifunctional code whereas the other rater assigned one function, were considered a match if the assigned single function was one of the functions subsumed under the multifunctional rating. Furthermore, raters agreed a revised definition of 'hedging' during the IRR process (see Appendix 2 for more details on this). Across both speaker groups, percentage agreements for form (87% for the CP group, 86% for the aphasic group) and syntax (96% for the CP group, 93% for the aphasic group) were higher than for function (70% for the CP group and 72% for the aphasic group). While IRR for form and syntax was higher than for function, all IRR figures were within an acceptable range, considering the IDK rating scheme was a novel instrument applied to aphasic discourse in dyadic exchanges (Kopenhaver Haidet, Tate, Divirgilio-Thomas, Kolanowski, & Happ, 2009).

We performed the Fisher's exact test to determine whether there was an association between speaker group and distributions of IDK forms, syntactic variations and functions.[4] We used Cramer's V as a measure of effect size (Field, 2009).[5] Adjusted residuals (z-scores) with a cut-off of +/-2 were used to identify which cells deviated from average values.

To illustrate functional patterns of IDK usage, representative examples of IDK functions were selected and analysed further, using existing conversation analytic (CA) transcriptions from Carragher et al. (2013) and Best et al. (2016). Appendix 3 includes information on CA transcription symbols used in these extracts.

---

[4] Due to frequencies below 5 in some cells, the chi-square test could not be applied.
[5] Conventions for Cramer's V with df = 3: .06 = small effect, .17 = medium effect, .29 = large effect; Cramer's V with df = 4: .05 = small effect, .15 = medium effect, .25 = large effect; Goss-Sampson, 2018.

Results

The following sections characterize the frequencies of IDK phonetic forms, syntactic variations and conversational functions separately for the two speaker groups.

What are IDK usage patterns in speakers with aphasia when engaged in dyadic conversation, and do these patterns differ between aphasic and CP groups?

The dataset yielded 62 CP IDK tokens and 77 IDK tokens produced by speakers with aphasia. All nine CPs contributed at least one IDK token (range: 1-9), whereas seven out of nine speakers in the aphasic group produced between one and 22 IDK tokens. Figure 2 shows the distribution patterns of IDK phonetic forms. The Fisher-exact test showed that the proportions of IDK forms differed across the two groups ($p = .032$, $V = .250$, medium effect). Inspecting the adjusted residuals of individual cells, this association was driven by the "*don't know*" (adjusted residuals: +/-2.8) and "*I don't know*" (adjusted residuals: +/-2.1) categories. Frequencies for "*don't know*" were considerably lower in the aphasic group (5% compared to 21% in the CP group) than what would be expected if counts were independent of speaker group. Frequencies in the "*I don't know*" category were considerably higher in the aphasic group (77% compared to 60% in the CP group). The least frequent form in the CP group was "*dunno*" (7%), which was also rarely produced by the aphasic group (8%). "*I dunno*" accounted for 13% of CP tokens and 10% of tokens produced by speakers with aphasia.

(Figure 2 near here)

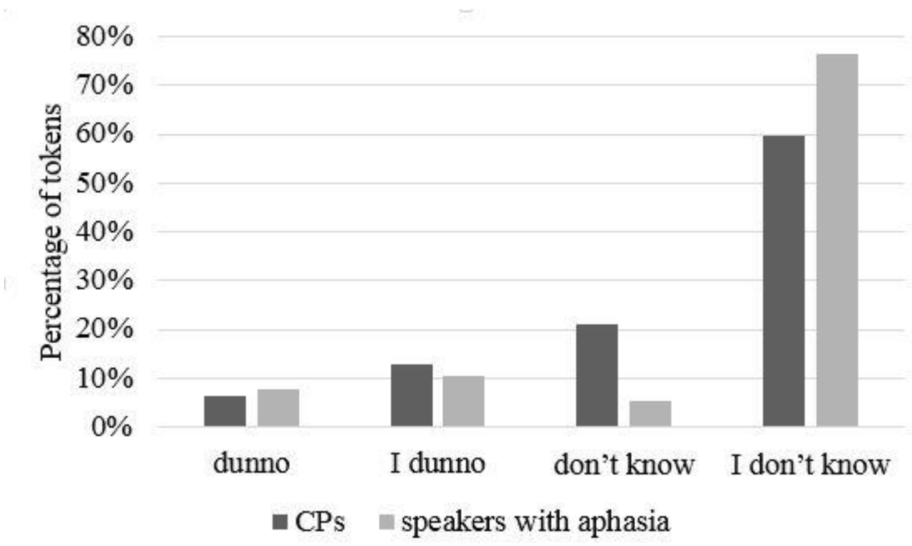

Figure 2. IDK phonetic forms in CPs and speakers with aphasia.

Figure 3 shows the distribution of syntactic variations. In order to calculate the Fisher's exact test, IDK tokens needed to be assigned to mutually exclusive categories. This was not the case when including counts for IDK + discourse marker, as these tokens could have co-occurred with other syntactic variations. Hence, we excluded this category from the analysis. Based on the remaining four syntactic variations, the distribution differed significantly between the two speaker groups ($p < .001$, Fisher's exact test, $V = .574$, large effect). Adjusted residuals indicated that this effect was largely driven by IDK tokens with complement (adjusted residuals: +/-6.1), but also by isolated IDK tokens and "other" (adjusted residuals: +/-3.6 and +/-3.7, respectively).

In the CP group, more than half of the tokens (35 out of 65; 54%) were IDK with complement, whereas speakers with aphasia rarely added a complement (6 out of 77; 8%), reflecting their language difficulties. Isolated IDK tokens were more common in the aphasic group (50/77 tokens; 65%) compared to the CP group (24/65 tokens; 37%). One speaker with aphasia, BL, made use of an unusual syntactic variation captured via the syntactic category "other" (18% of all tokens in the aphasic group): 14 out of 21 IDK tokens produced by BL represent the set phrase "*I don't know forget*" or "*I don't know for*". IDK tokens were accompanied by a discourse marker four times in the CP group (6%) and six times in the aphasic group (8%). In both groups, tokens with a *wh*-word were rarely produced (one token in the aphasic group; 1%, and two tokens in the CP group; 3%).

(Figure 3 near here)

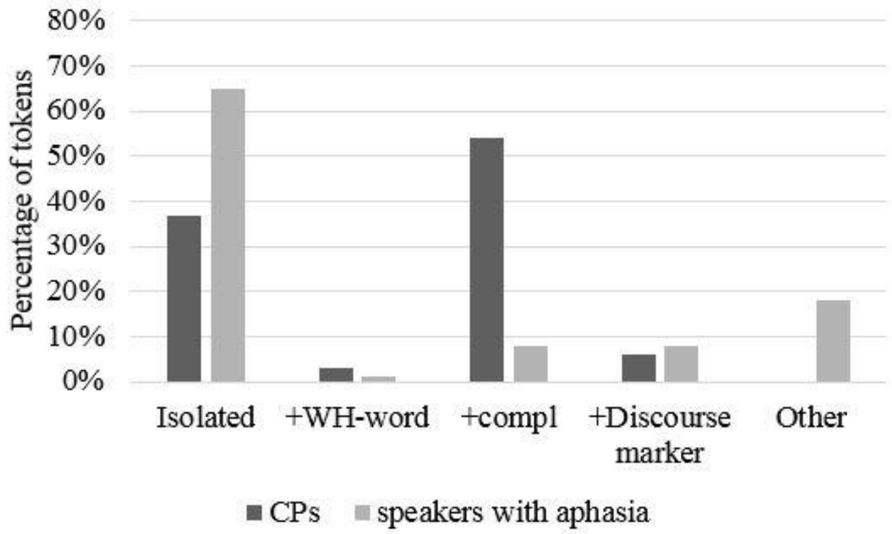

Figure 3. Syntactic variations of IDK tokens in CPs and speakers with aphasia.

Appendix 4 shows an overview of profiles of IDK functions for each individual. Figure 4 shows that IDK was used for a variety of functions by speakers with aphasia. However, there was a significant difference between the speaker groups with regard to the overall distribution of functions (p = .031, Fisher's exact test, V = .276, large effect). Inspection of adjusted residuals revealed that this effect was due to the lower number of multifunctional tokens in the aphasic group (4%; adjusted residual: -2.5) compared to the overall average (9%).

In both speaker groups, IDK frequently had a turn-constructional function (CPs: 32%, speakers with aphasia: 42%) or was used to indicate lack of knowledge (CPs: 31%, speakers with aphasia: 35%). The third most common function in both speaker groups was interpersonal (CPs: 21%, speakers with aphasia: 14%), for example avoiding commenting or hedging. There were a number of unclear IDK tokens in the aphasic group (5%), all produced by one speaker (BL).

(Figure 4 near here)

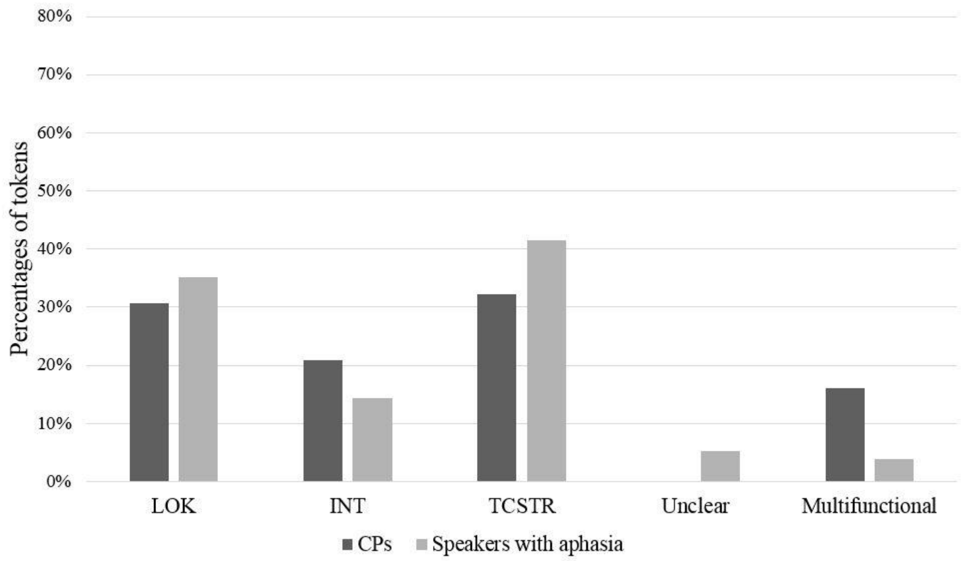

Figure 4. Functions of IDK tokens in CPs and speakers with aphasia.

A closer look at the relationship between syntactic variation and function revealed differences between speaker groups with regard to the functions of isolated IDK tokens. A large part (46%, 11/24) of CP isolated tokens were turn-constructional, and 13% (3/24) were coded as lack of knowledge. The remaining tokens were coded as either interpersonal (21%, 5/24) or multifunctional (21%, 5/24). In the aphasic group, on the other hand, isolated IDK tokens were used equally to indicate lack of knowledge (37%, 18/50), and to serve a turn-constructional function (37%, 17/50). The remaining isolated tokens were rated as interpersonal (22%, 11/50), multifunctional (4%, 2/50) and unclear (4%, 2/50). The higher percentage of isolated IDK tokens with lack of knowledge function in the aphasic group compared to the CP group might reflect a difficulty with constructing longer conversational turns. It could also be explained by the tendency of CPs to ask speakers with aphasia questions. In the aphasic group, 73% (19/27) of all IDK lack of knowledge tokens followed a CP question (and out of these, 14 were isolated IDK tokens, versus 32% (6/19) of CP tokens). While some examples suggest that these IDK tokens indicate an inability to answer a CP question due to aphasia ("*What's the name of the hotel, can you remember?*" – *"dunno"*, DM_4.2_1_2), others may show a more typical use of IDK, i.e., a genuine lack of knowledge ("*Have they got them in here?*" [referring to item in catalogue] – "*don't know",* Simon_2_3). In the CP group, similar lack of knowledge uses were found, for instance when a speaker with aphasia asks his CP "*[...] time is it?*" – "*I don't know*" (PM_1.3_1_2_3).

Within the turn-constructional category, both speaker groups used IDK tokens to hold or to yield the conversational floor. However, these functions were considerably more frequent in the aphasic group, in which these two sub-functions combined accounted for 81% of all TC tokens, compared to only 45% in the CP group. In the aphasic group, holding the conversational floor was often related to word finding difficulties, as in this example about refurbishing a conservatory: "*em em ((sings melody while gestures painting something)) em (0.3) I don't know em blue no em white no brown? brown? yeah.*" (JH_2.3_4). Extract 1 shows an example of a turn-yielding IDK from a speaker with aphasia. Here IDK is combined with a pause during which the speaker with aphasia looks at the CP (line 14) emphasizing the use of IDK to signal speaker change.

Extract 1 (JH_4.3_1)

	01	CP:	Tuesday (.) anticoagulant clinic
	02	PWA:	oh (for) flip sake (1.2) em, em, right so ⌈Northfields em (1.0) ⌉
	03		| ((draws circle on table with | ⌊finger, looks at CP)) ⌋
	04	CP:	mmm, ⌈mmm ⌉
	05	PWA:	⌊ ((looks away from CP)) ⌋
	06		⌈what (will) ⌉
	07		⌊((interrogative gesture with left hand, palm up)) ⌋ >em, em, em<
	08		⌈ (1.0) ⌉
	09		⌊((breathing out, lowers left arm with palm facing down)) ⌋
	10		right, ⌈or there ⌉
	11		⌊((raises left arm with palm facing down)) ⌋
	12		⌈(no) em, how much ⌉
	13		⌊((interrogative gesture with left hand, palm up)) ⌋
→	14		⌈em, em, **I don't know** ⌉ ⌈ (2.3) ⌉
	15		⌊((interrogative gesture with left hand, palm up)) ⌋ ⌊ ((looks at CP)) ⌋
	16	CP:	what, the reading?
	17	PWA:	Yes

CP examples of using IDK to hold and yield the conversational floor include utterances such as "*and if it's not raining it's sports day (1.4) but (0.9) I th- (0.7) I dunno (1.7) if it rains it* *could be off won't it, NAME'll let us know anyway*" (PM_3.2_8, where *"I dunno"* is followed by a lengthy lapse in the talk where either speaker could take a next turn).

Some speakers in the CP group showed a turn-constructional pattern that was not observed in the aphasic group; that of other-initiated repair (4/20 TC tokens). That is, the CP explicitly stated that he or she had trouble understanding the meaning of the speaker with aphasia's previous turn(s). Examples are "*I don't know what you want*" (CP; KK 2.2_2) or "*well I don't know what you mean love*" (CP; PG_4.2_5_6_7). This function was clearly related to expressive difficulties caused by aphasia.

Interpersonal and multifunctional IDK tokens were found in both speaker groups, however, they were less frequent in the aphasic group (14% and 4%, respectively, compared to 21% and 16% in the CP group). Extract 2 illustrates two IDK tokens with interpersonal functions used by a speaker with aphasia (one of which was coded as multifunctional). The conversation is about a mutual acquaintance who is attending car maintenance classes. The CP, at line 1, brings up the possibility of the speaker with aphasia attending a similar class.

Extract 2 (KK 3.2_1_2_3)

	01	CP:	you could go and do something like that
→	02	PWA:	hmm: no **I don't know** no hahahah
	03	CP:	be too hard cos of your speech?
→	04	PWA:	er: nothing (0.3) an (0.9) **I don't know**
	05		(1.1)
	06	CP:	°hmm° (.) bit more difficult though int °e?°

The first IDK at line 2, softens the speaker with aphasia's disagreement with this suggestion. This interpersonal function is emphasized by the accompanying laughter, a sign of a delicate issue or a dispreferred response (see Tsui, 1991). After the CP wonders whether such a class might be too hard, the speaker with aphasia makes a statement with a turn-final IDK (line 4). This was classified as multifunctional, as it a) serves as a turn-yielding device (the conversation lapses for 1.1 seconds afterwards), and b) signals avoidance of a commitment to the previous answer "no" (line 2).

A variety of functions (TC, M, LOK, and INT) were assigned to IDK + discourse marker combinations in both speaker groups. While the number of these items was low (6 tokens in aphasic group, 4 tokens in CP group), "*well*" was the most common co-occurring discourse marker (3 in the aphasic group and 2 in the CP group).

Finally, one speaker with aphasia (BL) made use of atypical IDK constructions, "*I don't know forget*" and "*I don't know for"*, sometimes combined with a gesture. These variations were observed in 67% of BL's IDK tokens (14/21). Two of these were assigned an "unclear" function, whereas the remaining 11 represented a turn-constructional function, four of which were used in order to take a turn. Only one other speaker with aphasia (KK) used IDK to take a turn, on one occasion. Extract 3 illustrates this idiosyncratic turn-taking function of IDK in BL's talk. Prior to this extract, BL pointed at the camera as the CP was leaving the room. At line 4, the CP sits back down on the couch.

Extract 3 (BL 1.8_8_9_10)

	01	CP:	OOH (0.5) it's cold when you move(d)
	02	PWA:	(ælə?)
	03		⌈ (1.4) ⌉
	04		⌊((CP sits back down)) ⌋
	05	CP:	what?
	06		(0.7)
→	07	PWA:	**I don't know for⌈ get** ⌉
	08		⌊((demonstrative gesture towards camera))⌋
	09		(2.5)
→	10	PWA:	⌈**I don't know forget** ⌉
	11		⌊ ((demonstrative gesture towards camera)) ⌋
	12		⌈ (2.5) ⌉
	13		⌊ ((CP folds her arms)) ⌋
	14	PWA:	⌈yes? ⌉
	15		⌊ ((looks at CP)) ⌋
	16		(0.5)
	17	CP:	°yeah what°?
	18		(0.6)
→	19	PWA:	⌈ **I don't ↑know**⌉ (2.7) camera, (0.3) shut, (0.9) the, (0.4) ⌈door. ⌉
	20		⌊((shakes head)) ⌋ ⌊((looks at CP))⌋
	21		⌈ (1.7) ⌉
	22		⌊ ((CP looks at PWA)) ⌋

BL produces an unintelligible turn at line 2 which is followed by a pause of 1.4 seconds before the CP indicates that she has trouble understanding, asking "*what?*" (line 5). This initiation of repair is followed by a pause and the two IDK tokens of interest here, at lines 7 and 10.

The IDK at line 7, combined with a demonstrative gesture towards the camera, appears to aid BL to take a turn. After a pause of 2.5 seconds, BL repeats his turn with identical intonation and gesture (line 10). Again, the IDK token appears to be a strategy to take a verbal turn, with semantic content added via the demonstrative gesture. It may be a comment about the fact that they were in the middle of making a video recording when the CP left the room. At line 14, after a significant pause, BL checks the CP's understanding with "*yes?*". However, the CP again initiates repair ("*yeah what*?", line 17), after which another turn-constructional IDK can be observed at line 19. This time the IDK token appears to hold the conversational floor, as the turn then continues with a comment about the camera. This reinforces the view that BL's IDK comments at lines 7 and 10 have been about the topic of video recording.

In summary, IDK was common to both speaker groups and available to all but two of the individuals with aphasia. The main difference between the two speaker groups was the proportion of isolated IDK tokens versus IDK tokens with complement. While instances of all function categories were found in neurotypical speakers as well as speakers with aphasia, our findings indicate group-specific usage patterns including a higher proportion of turn-constructional IDK tokens in the aphasic group. In addition, IDK was used as a turn-taking resource exclusively by two speakers with aphasia. By contrast, IDK as a resource for initiating repair was found to be a CP-specific function.

Discussion

We explored whether residual constructions in aphasic conversation consisted of high frequency, familiar word combinations. We analysed data from nine dyads recorded within everyday conversations, as well as a larger sample of semi-structured interviews with speakers with Broca's aphasia. Moreover, we presented usage patterns of IDK in everyday conversations of speakers with Broca's aphasia and their CPs.

Frequency-based analysis

Our frequency-based analysis showed that association measures such as t-scores are an effective way of quantifying well-formed aphasic language output. They are robust across conversational settings (everyday conversations versus semi-structured interview data) and across individuals with varying degrees of aphasic impairment. Collocation strength as measured by t-scores can be used as an estimate of the degree of formulaicity (Zimmerer et al., 2018), with higher scores indicating a higher likelihood that a word combination represents (part of) a formulaic expression. We observed more strongly collocated combinations in the aphasic group compared to neurotypical CPs – a finding that is consistent with previous research (Zimmerer et al., 2018), and, if seen as indicators of formulaic language, also consistent with studies using different methods (Van Lancker Sidtis & Postman, 2006). Our study shows that increased reliance on familiar collocations in aphasia also extends to everyday conversational settings. Another contribution is the analysis of both types and tokens as well as bigrams and trigrams. At the level of bigrams, type- but not token-based inventories distinguish the aphasic group from CPs. However, at the level of trigrams, groups were distinguished based on both types and tokens. Effect sizes were large for all type- and token-based comparisons both at the level of bi- and trigrams, and the preserved inventory of common phrases in aphasia appears to be influenced by usage-based factors such as frequency and collocation strength (DeDe, 2013; Knilans &

DeDe, 2015). These strongly collocated residual utterances require less combinatorial effort and may be processed as holistic units (Zimmerer et al., 2018), which makes them resilient to aphasia, particularly where combinatorial mechanisms are disrupted.

The stability of frequency-based profiles across different talk contexts suggests ecological validity of elicited speech tasks when taking association measures. Such tasks represent more controlled settings compared to everyday conversations and allow easier comparisons between individuals/groups as there is less variability as to the content/nature of the interaction. Future investigations could use association measures in tasks such as narrative production for comparisons across individuals or speaker groups, or to investigate further the influence of aphasia severity on the reliance on familiar collocations.

It should be noted that frequency of use is a complex variable, intercorrelated with other measures such as age-of-acquisition (AoA; Baayen, Milin, & Ramscar, 2016). Just as AoA may help to explain why some single words remain accessible despite aphasia (Brysbaert & Ellis, 2016), it could be a confounding variable of frequency-related effects at the multiword level. A recent study with neurotypical participants provided evidence of multiword AoA effects on language processing (Arnon, McCauley, & Christiansen, 2017). Future studies addressing the processing of familiar collocations in aphasia might consider AoA as well as association measures.

Zimmerer et al. (2018) listed the 10 most frequent trigram types produced by speakers with non-fluent aphasia. Interestingly, four of these overlap with the most frequently produced trigrams in the present naturalistic dataset (Table 4): "*don't know*", "*I don't*", trigrams with numerals ("*one or two*" / "*one two three*" / "*two three four*"), and "*it's X*" ("*it's a*" / "*it's alright*"). All of these represent strongly collocated, relatively simple clausal structures. Previous studies have shown that non-fluent aphasia affects complexity of spoken output (e.g., more accurate production of simple, as compared to complex verbs; Thompson et al., 1997). It would be interesting to analyse the internal structure / complexity of familiar phrases to explore other factors that underpin presence of constructions in aphasia. To date, this has not been investigated from a constructivist, usage-based perspective. Another area for further research is an exploration of ungrammatical utterances, examining the frequency and nature of errors in strings that differ in degree of formulaicity.

IDK analysis

Despite this frequency-based difference between the aphasic and CP group, and the variability in the types of word combinations accessible to speakers with aphasia, one shared construction, "*don't know*", was found in the inventories of both speaker groups. IDK stands out in Broca's aphasia because its syntactic structure is atypical of agrammatic output. However, and in contrast to studies of automatic speech, recurrent or stereotyped phrases that imply this is pathological language (Blanken & Marini, 1997; Grande et al., 2008; Rodrigues & Castro-Caldas, 2014), our results suggest that familiar collocations such as IDK may be 'stereotypes' at a formal, but not at a functional level.

IDK was most commonly realized as the full form "*I don't know*" in the aphasic and control group. There was a small difference of overall distribution of IDK phonetic forms across the two groups, however, a larger effect size was seen when comparing the group level syntactic variations of IDK tokens. While speakers with aphasia most often used isolated IDK tokens, CPs produced more IDK tokens with a complement. This finding was expected since speakers with aphasia have difficulties in combining smaller linguistic units into longer, grammatically well-formed utterances. At the same time, this suggests that IDK is represented as a relatively fixed unit that may not require grammatical processing, a finding that supports the claim that familiar collocations like IDK may be processed as formulas. On the other hand, combinations of IDK and discourse markers such as "*well*" could be observed in the aphasic group with similar frequency to the CP group, indicating that combining IDK with pragmatic elements may be easier than adding lexical elements. However, the overall number of these cases was small.

IDK fulfilled a variety of functions in both CPs and speakers with aphasia. When comparing the frequencies of IDK functions across the two speaker groups, the findings reveal a group difference with a large effect size. This difference stems from the relatively low number of multifunctional IDK tokens in the aphasic group. However, this result needs to be interpreted with caution given that IRR for function was lower than for phonetic form and syntax. The function most commonly associated with IDK, namely indicating a lack of knowledge, was observed in both CPs and speakers with aphasia. However, the number of isolated IDK tokens with a lack of knowledge function was higher in the aphasic group compared to CPs. In our dataset, question-answer sequences initiated by a CP were common, and this may account for the high number of IDK tokens produced by speakers with aphasia that served to signal lack of knowledge. Both the amount of questioning and the response type reflects the presence of aphasia; CPs use questions to initiate conversation with a person with aphasia, and speakers with aphasia use IDK to provide legitimate answers to such questions whilst signalling the presence of aphasic language difficulties.

IDK was frequently used by speakers with aphasia for turn-constructional functions such as turn yielding, a finding supported by Simmons-Mackie & Damico (1997). Turn-constructional functions were also found in CPs which is in line with previous studies with neurotypical speakers (Pekarek Doehler, 2016). The higher proportion of tokens assigned to turn holding/yielding in the aphasic group is likely to stem from difficulties associated with aphasia. A speaker with aphasia may rely on the use of an island of fluency such as IDK as a resource for opening the conversational floor when grammatical and word finding difficulties make turn construction challenging, or as in extract 3, to regulate turn-taking. The use of familiar collocations to aid turn construction has also been previously reported in Broca's aphasia (Beeke, 2003). Another finding directly related to aphasia was a CP-specific pattern, namely IDK as a method of initiating repair, i.e. signalling a lack of understanding of a prior aphasic turn.

In addition, some speakers with aphasia used IDK for interpersonal functions such as to avoid commenting and, in a small number of cases, IDK turns could even be described as multifunctional (e.g., with mixed turn-constructional and interpersonal functions). Despite the relatively low frequency of such tokens, such multiple functions – associated with typical IDK usage – have not been documented in speakers with aphasia to date. Hence, the present study suggests that IDK is a conversational building block in Broca's aphasia that extends beyond the reported turn-constructional function (Simmons-Mackie & Damico, 1997) and the more fundamental lack of knowledge function. This suggests that such building blocks may be invaluable linguistic structures that could routinely be considered in language assessments and could be harnessed in speech and language therapy interventions (e.g., Stahl & Van Lancker Sidtis, 2015).

Conclusion

The observation of more strongly collocated word combinations in speakers with aphasia supports a usage-based view of language processing. Just as many speakers with aphasia are able to retrieve familiar single words, they can also retrieve familiar multi-word utterances such as IDK.

This study suggests that identifying common word combinations in aphasic talk, with help of a frequency-based approach, is useful to characterize and evaluate the grammatical behaviour of individuals with aphasia. IDK was used in different communicative situations as a relatively fixed, yet effective conversational tool with a functional profile that seems to be adjusted to the turn construction difficulties associated with aphasia.

Authorship statement and acknowledgements

SB and MC acquired the naturalistic data. CB analysed the data, under the supervision of SB, RV and VZ. GB was rater 2 in the IDK analysis. CB wrote the first draft of the manuscript. SB, RV, VZ, MC and GB contributed to further drafts. We thank all participants who provided everyday conversational data, and the AphasiaBank team for access to semi-structured interview data. We also thank student researchers who transcribed some of the video samples, and two anonymous reviewers for their valuable comments.

This work was supported by an Economic and Social Research Council Training Grant (1475581) awarded to the first author.

Disclosure of interest

The authors have no financial conflicts of interest.

References

Arnon, I., McCauley, S. M., & Christiansen, M. H. (2017). Digging up the building blocks of language: Age-of-acquisition effects for multiword phrases. *Journal of Memory and Language*, *92*, 265–280. http://doi.org/10.1016/j.jml.2016.07.004

Baayen, R. H., Milin, P., & Ramscar, M. (2016). Frequency in lexical processing.

Aphasiology, 30(11), 1174–1220. http://doi.org/10.1080/02687038.2016.1147767

Beeke, S. (2003). “I suppose” as a resource for the construction of turns at talk in agrammatic aphasia. *Clinical Linguistics & Phonetics, 17*(4–5), 291–298. http://doi.org/10.1080/0269920031000080055

Beeke, S., Maxim, J., Best, W., & Cooper, F. (2011). Redesigning therapy for agrammatism: Initial findings from the ongoing evaluation of a conversation-based intervention study. *Journal of Neurolinguistics, 24*(2), 222–236. http://doi.org/10.1016/j.jneuroling.2010.03.002

Best, W., Maxim, J., Heilemann, C., Beckley, F., Johnson, F., Edwards, S. I., … Beeke, S. (2016). Conversation Therapy with People with Aphasia and Conversation Partners using Video Feedback: A Group and Case Series Investigation of Changes in Interaction. *Frontiers in Human Neuroscience, 10*(562). http://doi.org/10.3389/fnhum.2016.00562

Blanken, G. (1991). The functional basis of speech automatisms (recurring utterances). *Aphasiology, 5*(2), 103–127. http://doi.org/10.1080/02687039108249477

Blanken, G., & Marini, V. (1997). Where do lexical speech automatisms come from? *Journal of Neurolinguistics, 10*(1), 19–31. http://doi.org/10.1016/S0911-6044(97)00001-8

Brysbaert, M., & Ellis, A. W. (2016). Aphasia and age of acquisition: are early-learned words more resilient? *Aphasiology, 30*(11), 1240–1263. http://doi.org/10.1080/02687038.2015.1106439

Carragher, M., Sage, K., & Conroy, P. (2013). The effects of verb retrieval therapy for people with non-fluent aphasia: evidence from assessment tasks and conversation. *Neuropsychological Rehabilitation*, *23*(6), 846–87. http://doi.org/10.1080/09602011.2013.832335

Carragher, M., Sage, K., & Conroy, P. (2015). Outcomes of treatment targeting syntax production in people with Broca's-type aphasia: evidence from psycholinguistic assessment tasks and everyday conversation. *International Journal of Language & Communication Disorders*, *50*(3), 322–336. http://doi.org/10.1111/1460-6984.12135

Code, C. (1982). Neurolinguistic analysis of recurrent utterances in aphasia. *Cortex*, *18*, 141–152.

Conrad, S., & Biber, D. (2004). The Frequency and Use of Lexical Bundles in Conversation and Academic Prose. *Lexicographica*, *20*, 56–71.

DeDe, G. (2013). Verb transitivity bias affects on-line sentence reading in people with aphasia. *Aphasiology*, *27*(3), 326–343. http://doi.org/10.1080/02687038.2013.843151

Diani, G. (2004). The discourse functions of I don't know in English conversation. In K. Aijmer & A.-B. Strenstrom (Eds.), *Discourse Patterns in Spoken and Written Corpora* (pp. 157–171). John Benjamins.

Drew, P., & Holt, E. (1998). Figures of speech: Figurative expressions and the management of topic transition in conversation. *Language in Society*, *27*(4), 495–522. http://doi.org/10.1017/S0047404598004035

Druks, J., & Masterson, J. (2000). *An object and action naming battery*. Psychology Press.

Durrant, P., & Doherty, A. (2010). Are high-frequency collocations psychologically real? Investigating the thesis of collocational priming. *Corpus Linguistics and Linguistic Theory*, *6*(2), 125–155. http://doi.org/10.1515/CLLT.2010.006

Erman, B., & Warren, B. (2000). The idiom principle and the open choice principle. *Text - Interdisciplinary Journal for the Study of Discourse*, *20*(1), 29–62. http://doi.org/10.1515/text.1.2000.20.1.29

Field, A. (2009). *Discovering statistics using SPSS* (Third edit). SAGE Publications.

Gahl, S., & Menn, L. (2016). Usage-based approaches to aphasia. *Aphasiology, 30(11)*, 1361-1377. http://doi.org/10.1080/02687038.2016.1140120

Goldberg, A. E. (2003). Constructions: a new theoretical approach to language. *Trends in Cognitive Sciences*, *7*(5), 219–224. http://doi.org/10.1016/S1364-6613(03)00080-9

Goss-Sampson, M. A. (2018). Statistical analysis in JASP: A students guide. Retrieved from https://jasp-stats.org/jasp-materials/

Grande, M., Hussmann, K., Bay, E., Christoph, S., Piefke, M., Willmes, K., & Huber, W. (2008). Basic parameters of spontaneous speech as a sensitive method for measuring change during the course of aphasia. *International Journal of Language & Communication Disorders*, *43*(4), 408–26. http://doi.org/10.1080/13682820701685991

Grant, L. E. (2010). A corpus comparison of the use of I don't know by British and New Zealand speakers. *Journal of Pragmatics*, *42*(8), 2282–2296. http://doi.org/10.1016/j.pragma.2010.01.004

Gries, S. T. (2010). Useful statistics for corpus linguistics. In A. Sánchez & M. Almela (Eds.), *A mosaic of corpus linguistics: selected approaches* (pp. 269–291). Frankfurt am Main: Peter Lang.

Hesson, A. M., & Pichler, H. (2016). Interpreting "I don't know" use by persons living with dementia in Mini-Mental State Examinations. *Patient Education and Counseling*, *99*(9), 1534–1541. http://doi.org/10.1016/j.pec.2016.07.001

Hildebrand-Edgar, N. (2016). *Disentangling frequency effects and grammaticalization* (Working Papers of the Linguistics Circle of the University of Victoria No. 26,1). Retrieved from https://journals.uvic.ca/index.php/WPLC/article/view/15301

Hunston, S. (2002). *Corpora in applied linguistics*. Cambridge: Cambridge University Press.

Kaplan, E., Goodglass, H., & Weintraub, S. (2001). *Boston Naming Test. 2nd Ed.* Philadelphia, PA: Lippincott Williams & Wilkins.

Kecskés, I. (2000). A cognitive-pragmatic approach to situation-bound utterances. *Journal of Pragmatics*, *32*(5), 605–625. http://doi.org/10.1016/S0378-2166(99)00063-6

Kertesz, A. (2007). *Western Aphasia Battery-Revised*. San Antonio: PsychCorp.

Knilans, J., & DeDe, G. (2015). Online sentence reading in people with aphasia: evidence from eye-tracking. *American Journal of Speech-Language Pathology, 24*(4), S961–S973. http://doi.org/10.1044/2015_AJSLP-14-0140

Kopenhaver Haidet, K., Tate, J., Divirgilio-Thomas, D., Kolanowski, A., & Happ, M.B. (2009). Methods to improve reliability of video recorded behavioural data. *Research in Nursing & Health, 32*(4), 465-474. doi: 10.1002/nur.20334

Lum, C., & Ellis, A. W. (1999). Why do some aphasics show an advantage on some tests of nonpropositional (automatic) speech? *Brain and Language*, *70*(1), 95–118. http://doi.org/10.1006/brln.1999.2147

MacWhinney, B., Fromm, D., Forbes, M., & Holland, A. (2011). AphasiaBank: Methods for studying discourse. *Aphasiology, 25*(11), 1286–1307. doi:10.1080/02687038.2011.589893

Meunier, F. (2012). Formulaic language and language teaching. *Annual Review of Applied Linguistics*, *32*, 111–129. http://doi.org/10.1017/S0267190512000128

Mikesell, L. (2009). Conversational Practices of a Frontotemporal Dementia Patient and His Interlocutors. *Research on Language and Social Interaction*, *42*(2), 135–162. http://doi.org/10.1080/08351810902864552

Pekarek Doehler, S. (2016). More than an epistemic hedge: French je sais pas “I don”t know’ as a resource for the sequential organization of turns and actions. *Journal of Pragmatics*, *106*, 148–162. http://doi.org/10.1016/j.pragma.2016.06.014

Pichler, H., & Hesson, A. (2016). Discourse-pragmatic variation across situations, varieties, ages: I DON’T KNOW in sociolinguistic and medical interviews. *Language and Communication*, *49*, 1–18. http://doi.org/10.1016/j.langcom.2016.04.002

Rodrigues, I. T., & Castro-Caldas, A. (2014). Aphasia with recurring utterances: Old syndrome, new perspectives. *Aphasiology*, *28*(11), 1350–1363. http://doi.org/10.1080/02687038.2014.921884

Schmitt, N., & Carter, R. (2004). Formulaic Sequences in Action: An Introduction. In N. Schmitt (Ed.), *Formulaic Sequences: Acquisition, Processing, and Use* (pp. 1–22). Amsterdam, the Netherlands: John Benjamins.

Sidtis, D., Canterucci, G., & Katsnelson, D. (2009). Effects of neurological damage on production of formulaic language. *Clinical Linguistics & Phonetics*, *23*(4), 270–84. http://doi.org/10.1080/02699200802673242

Simmons-Mackie, N. N., & Damico, J. S. (1997). Reformulating the definition of compensatory strategies in aphasia. *Aphasiology*, *11*(8), 761–781. http://doi.org/10.1080/02687039708250455

Stahl, B., & Van Lancker Sidtis, D. (2015). Tapping into neural resources of communication: formulaic language in aphasia therapy. *Frontiers in Psychology*, *6*, 1526. http://doi.org/10.3389/fpsyg.2015.01526

The British National Corpus, version 3 (BNC XML Edition). (2007). Distributed by Oxford University Computing Services on behalf of the BNC Consortium. Retrieved from http://www.natcorp.ox.ac.uk/

Thompson , C.K., Lange, K.L., Schneider, S.L., & Shapiro, L.P. (1997). Agrammatic and non-brain-damaged subjects' verb and verb argument structure production. *Aphasiology, 11*(4-5), 473-490. https://doi.org/10.1080/02687039708248485

Tsui, A. B. M. (1991). The pragmatic functions of I don't know. *Text & Talk*, *11*(4), 607–622. http://doi.org/10.1515/text.1.1991.11.4.607

Van Lancker, D., & Kempler, D. (1987). Comprehension of familiar phrases by left- but not by right-hemisphere damaged patients. *Brain and Language*, *32*(2), 265–277. http://doi.org/10.1016/0093-934X(87)90128-3

Van Lancker Sidtis, D., & Postman, W. A. (2006). Formulaic expressions in spontaneous speech of left- and right-hemisphere-damaged subjects. *Aphasiology*, *20*(5), 411–426. http://doi.org/10.1080/02687030500538148

Van Lancker Sidtis, D., & Rallon, G. (2004). Tracking the incidence of formulaic expressions in everyday speech: methods for classification and verification. *Language & Communication*, *24*(3), 207–240. http://doi.org/10.1016/j.langcom.2004.02.003

Wray, A. (2002). The Whole and the Parts. In A. Wray (Ed.), *Formulaic Language and the Lexicon* (pp. 3–18). Cambridge: Cambridge University Press.

Zimmerer, V. C., Newman, L., Thomson, R., Coleman, M., & Varley, R. (2018). Automated analysis of language production in aphasia and right hemisphere damage: frequency and collocation strength. *Aphasiology*, *32*(11), 1267–1283. http://doi.org/10.1080/02687038.2018.1497138

Zimmerer, V. C., & Wibrow, M. (2015). Frequency in Language Analysis Tool.

Zimmerer, V. C., Wibrow, M., & Varley, R. A. (2016). Formulaic language in people with probable Alzheimer's Disease: a frequency-based approach. *Journal of Alzheimer's Disease*, *53*(3), 1145–1160. http://doi.org/10.3233/JAD-160099