# FORMULAIC SEQUENCES AS FLUENCY DEVICES IN THE ORAL PRODUCTION OF NATIVE SPEAKERS OF POLISH

*EWA GUZ*
John Paul II Catholic University of Lublin
ewasik@o2.pl

**Abstract**
In this paper we attempt to determine the nature and strength of the relationship between the use of formulaic sequences and productive fluency of native speakers of Polish. In particular, we seek to validate the claim that speech characterized by a higher incidence of formulaic sequences is produced more rapidly and with fewer hesitation phenomena. The analysis is based on monologic speeches delivered by 45 speakers of L1 Polish. The data include both the recordings and their transcriptions annotated for a number of objective fluency measures. In the first part of the study the total of formulaic sequences is established for each sample. This is followed by determining a set of temporal measures of the speakers' output (speech rate, articulation rate, mean length of runs, mean length of pauses, phonation time ratio). The study provides some preliminary evidence of the fluency-enhancing role of formulaic language. Our results show that the use of formulaic sequences is positively and significantly correlated with speech rate, mean length of runs and phonation time ratio. This suggests that a higher concentration of formulaic material in output is associated with faster speed of speech, longer stretches of speech between pauses and an increased amount of time filled with speech.

**Keywords**: formulaic sequences, speed fluency, breakdown fluency, temporal speech measures

## 1. Introduction

Research into L1 speech suggests that formulaic language contributes significantly to fluent language production (Wray and Perkins 2000; Wray 2002; Wray 2008; Pawley 2009). Pawley and Syder (1983, 2000) report that native-like oral fluency depends largely on the speaker's ability to rely on automatized repertoires of prefabricated chunks which reduce the amount of processing and encoding involved in speech production and afford the speaker the time to attend to other aspects of the speaking process. Selecting a formulaic rather than word-by-word formulation of the utterance results in the preservation of the speaker's cognitive resources and has a direct bearing on the temporal characteristics of the utterances, which are produced rapidly and smoothly without hesitations and pauses (Raupach 1980; Gatbonton and Segalowitz 1988; Towell, Hawkins and Bazergui 1996). The central goal of the present study is to

establish whether the use of formulaic sequences in native speech is linked to the speaker's overall productive fluency. In particular, we attempt to determine the correlation between the number of formulaic sequences identified in the oral output of native speakers of Polish and a range of preselected speed and breakdown fluency measures. Our discussion begins with a brief definition of the two key concepts investigated in this study: fluency and formulaicity. This is followed by an overview of research into the relationship between fluency and the use of formulaic language in L1 and L2 speech. Finally, the research methodology adopted for the purposes of the study as well its results are described.

## 2. Fluency as a production/performance phenomenon

Speech fluency is a complex phenomenon which has been the subject of extensive research over the years. Earlier accounts of fluency saw it as a multifaceted construct which embraces a variety of aspects of language use ranging from sociolinguistic appropriateness through linguistic creativity to expressing ideas at length in coherent semantically dense sentences with few pauses and fillers (Fillmore 1979: 93). Later accounts define fluency as a psycholinguistic and speech-related phenomenon (Lennon 2000; Gatbonton and Segalowitz 2005) which has practical applications as a performance descriptor and indicator of progress (Chambers 1997). Understood in this way, fluency is synonymous with speech fluidity defined as the ability to efficiently "translate thought and communicative intentions into speech under the temporal constraints of online speech processing (Lennon 2000: 26)." Other accounts look at fluency from a more interactive perspective and postulate its additional 'perceived' or 'perceptive' dimension (Guillot 1999; Freed 2000; Kormos and Denes 2004; Segalowitz 2010; Götz 2013). As the main focus of this paper is speech production rather than reception, the definition of fluency adopted in this study corresponds to that of Segalowitz's (2010: 48) utterance fluency which he equates with the actual, quantifiable features of oral utterances gauged by a range of temporal measures reflecting the speed of speaking, time filled with speech vs silence, occurrence of pausing as well as hesitation and repair phenomena (for an extensive review of temporal measures used in fluency research see Kormos 2006). In fact, the need for adopting a range of clear-cut variables representing distinct aspects of speech is stressed by Skehan (2003), who recommends that any analysis of fluency should involve a separate treatment of such speech features as silence, speed of speaking, reformulation and automatization. In an attempt to provide a comprehensive and valid fluency description that avoids collapsing its distinctive temporal aspects into one dimension, Tavakoli and Skehan (2005) propose a division of fluency into speed, breakdown and repair fluency. In this framework, fluency is defined in purely temporal terms as a capacity to organize speech in real time and is captured by speed and breakdown fluency (Skehan 2003). The former denotes the number of words uttered per time unit, whereas the latter reflects time filled with speech and is expressed in measurements based on the number, length and location of pauses. A third, additional measure – repair fluency – corresponds to the number of false starts, misformulations, self-corrections and repetitions.

Taking into account the rationale underlying this study, which states that the use of formulaic language lessens the cognitive processing effort involved in online speech production, we have decided to adopt the approach to fluency proposed by Tavakoli and Skehan (2005) as one that captures accurately and reflects directly the temporal nature of speech production[1]. Basing on the calculation of the time spent speaking and pausing respectively, speed and breakdown fluency measures are indicative of the amount of effort needed by the speaker to produce utterances under real time constraints.

## 3. Definition of formulaic sequences

Recent years have seen a surge of scholarly interest in formulaic language in general and its involvement in the production of fluent L1 and L2 speech in particular (Wray 2002, 2008; Schmitt 2004; Meunier and Granger 2008ab; Corrigan et al. 2009ab; Wood 2001, 2006, 2007, 2010ab; Weinert 2010).

Although there seems to be a general consensus as to the reality of formulaic sequences, they are very difficult to define precisely. Despite the growing body of research in this area, there does not appear to be a single, all-inclusive definition of a formulaic sequence that would embrace the phenomenon in its entirety. Most researchers involved in investigating formulaic sequences emphasize that they constitute a notoriously heterogenous category which accommodates a variety of multiword strings marked by a varied degree of semantic, syntactic, phonological and pragmatic integrity (see for example Wray and Perkins 2000). Wray (2002: 9) provides a definition which has become widely recognized in formulaicity research and is gradually entering mainstream use. She defines a formulaic sequence as:

> (a) sequence, continuous or discontinuous of words or other elements, which is, or appears to be, prefabricated: that is stored and retrieved whole from memory at the time of use, rather than being subject to generation or analysis by the language grammar.

It is worth noting that Wray's definition clearly prioritizes the holistic storage and retrieval of formulaic language as its prototypical feature. In fact, the psycholinguistic unity of formulaic sequences has been a recurrent theme in the majority of studies. More specifically, one observation that has been frequently made across studies with reference to the single-word-like status of formulaic sequences is that they contribute to the speaker's overall fluency by reducing the cognitive pressures involved in the construction of fluent speech. This line of thinking is based on two key assumptions concerning formulaic sequences on the one hand and the nature of speech processing on the other. These will be discussed in section 4.

---

[1] Repair fluency is excluded form the scope of the analysis as reformulation phenomena were scarce and appeared only in some of the examined samples.

## 4. Formulaic sequences as fluency-devices

The central concern of the present study is to investigate the relationship between productive fluency and the use of formulaic language in L1with the view to determining whether the use of formulaic sequences frees up cognitive resources for attending to other aspects of speech, which in turn may lead to an overall more fluent performance in terms of speed of speech and pausing.

The first premise which underlies our reasoning is that spontaneous speech is largely made up of prefabricated word sequences which constitute the preferred lexical choices of native speakers of a given language within a particular speech community (see Peters 1983; Erman and Warren 2000; Schmitt 2004; Wray 2002, 2008). This aspect of natural language use, referred to as 'native-like selection' (Pawley and Syder 1983), expresses the tendency of native language users to rely on a relatively closed set of ready-made prefabricated multiword strings rather than draw on freely from a pool of possible combinations of single words bound by the rules of grammar alone. Such 'productive speech formulas' constitute the major building blocks of speech and are memorized and holistically retrieved from long-term memory (Pawley 2009). Numerous scholars investigating the patterning of naturally occurring language (Altenberg 1990; Sinclair 1991; Renouf and Sinclair 1991; Aijmer 1996; Eeg-Olofsson and Altenberg 1996; Cowie 1998; Moon 1998; Biber et al. 1999; Hunston and Francis 2000) emphasize the lexically-driven character of spoken language and argue in favour of a view of language which accommodates the ability of native speakers to store, retrieve and recognize formulaic sequences. In such models, speech production is seen as based on two modes of processing; analytic grammar-based generation (and generalisation) as opposed to holistic, item-based handling of formulaic language.

The second dimension of speech which lies at the core of the present discussion is that L1 speech production is a complex mental and highly automatized process that occurs in the presence of heavy time and memory capacity constraints. Kormos (2006: 38) describes human speech as an effortless skill whose mastery is underpinned by automaticity. In psycholinguistic accounts of speech, automaticity is defined as a set of "utterance production skills which refer to the speed and ease of handling utterances (Gabonton and Segalowitz 1988: 473)." Apart from automaticity, production of rapid speech depends on the process of proceduralization whereby declarative (linguistic) knowledge is converted into procedural one (Levelt 1989; Anderson 1983). In this perspective, the cognitive advantage of formulaic sequences comes from their automatized/proceduralized nature, which allows speakers to exceed the limitations imposed by the size of their memory. Handled as single words, the automatized/proceduralized formulaic strings are retrieved directly from long-term memory (without conscious effort, attention or control) bypassing the storage limit of short-term memory and undergo processing in the working memory in larger units (Towell, Hawkins and Bazergui 1996; Lennon 2000; Segalowitz 2003). This, in turn, allows speakers to produce speech in a more fluent manner i.e. with fewer pauses and longer fluent runs between pauses.

Taken together, the findings presented so far suggest that the use of formulaic sequences aids the production of native speech in that it saves the cognitive resources and time, which can be used to attend to other aspects of speech such as e.g. discourse

planning or manipulating information (Nattinger and DeCarrico 1992; Wray and Perkins 2000).

Although the validity of the relationship between fluency and formulaicity has been widely recognized, very few studies have investigated the exact nature of this association and little empirical evidence has been provided so far to validate its existence. Raupach (1984) is one of the first studies to emphasize the link between formulaicity and fluency. His study presents evidence of over-reliance on certain formulas (modifiers and rhetorical organizers) in German learners of L2 French. Raupach reports that his participants used formulaic sequences as a production strategy, or as 'islands of reliability.' Similarly, Dechert's (1984) investigation of German students of English (in a task which required them to retell a narrative in L2) reveals that the most fluent learners relied on a number of multiword sequences around which they pieced together a spoken narrative.

More recent accounts have focused on the significance of formulaic language in L2 fluency development. One such example is a series of longitudinal case studies conducted by David Wood (2004, 2006, 2009, 2010b) offering a fine grained analysis of the facilitating influence of automatized lexical phrases on the development of L2 temporal fluency. Wood (2004, 2006, 2010b) investigates fluency gains in individual ESL learners over an extended period of time and reports evidence of formulaic sequences functioning as fluency-enhancing devices. Wood (2009) presents the results of a longitudinal investigation of effects of focused instruction of formulaic sequences and fluency on the oral performance of L2 speakers. His results indicate a considerable improvement in temporal fluency after six weeks of focused instruction and a relationship between fluency and use of formulaic sequences in the learner speech samples. In a study based on a similar setup, De Jong, Halderman and Ross (2009) look at the effect of formulaic language instruction on speed and pause fluency in EFL learners at various levels of proficiency. On the one hand, their results show that an increase in the use of formulaic sequences leads to longer fluent runs. On the other hand, however, a higher number of formulaic strings in learner output is associated with an increase in the length of pauses. On the basis of these inconclusive results, the authors state that the use of formulaic sequences by L2 learners is not effortless and that the trained formulaic sequences were probably not stored as chunks and their retrieval was not automatized.

As regards investigating the fluency-enhancing role of formulaic language in L1 speech, to the knowledge of the author of the present investigation, there have been no studies to date that provide solid empirical evidence confirming this relationship between temporal fluency and the use of formulaic language. This study is an attempt at filling this gap.


## 5. The study

Basing on the assumption that formulaic sequences play some role in the speaker's productive fluency, our primary concern here is to investigate the relationship between the two aspects of native speech. In the analysis that follows, we set out to address the following research questions:

1.  Is there a relationship between the use of formulaic sequences and temporal fluency in the speech of Polish native speakers?
2.  Is there a correlation between the use of formulaic sequences and speed fluency as measured by articulation rate?
3.  Is there a correlation between the use of formulaic sequences and breakdown fluency as measured by the mean length of runs, mean length of pauses and phonation time ratio?
4.  Is there a correlation between the use of formulaic sequences and speech rate?

## 5.1 Participants

Participants included 45 undergraduate students (29 female, 16 male) at the University of Warsaw enrolled at the second year of a teacher training college and working towards their BA in teaching English as a Foreign Language. All of them were of Polish origin and spoke Polish as their first language. Prior to the study, participants were given extensive information about the character and purpose of the research. Written permissions were obtained to indicate explicit consent to participate in the study.

## 5.2 Procedure

The analysis is based on a 9,979 word data set consisting of L1 monologic speeches elicited through a personalising task (see Appendix 1 for a complete list of topics used as prompts). The data include both the recordings and their transcriptions annotated for pauses. In the first part of the study a set of temporal fluency measures of the speakers' output are determined. This is followed by establishing the total of formulaic sequences for each sample. Finally, to gauge the nature of the relationship between formulaicity and speech fluency Pearson product-moment coefficients are established for these two variables.

### 5.2.1 Measuring utterance fluency

As regards the pause annotation procedure, we followed a well-established research tradition (see Goldman-Eisler 1968; Towell, Hawkins and Bazergui 1996; Freed et al. 2004; Segalowitz and Freed 2004; Wood 2006; de Jong et al. 2012), where a pause is defined as a silence or a non-verbal filler of 0.25 seconds or longer. A visual representation was produced for each sample using PRAAT software (Boersma and Weenink 2005) to indicate pause boundaries and identify their duration. Pauses were then indicated in text using square brackets. Syllables were counted consulting the original recording when necessary. For reasons already explained, in this study we adopt the fluency measurement paradigm put forward by Skehan (2003, 2009) and Tavakoli and Skehan (2005) followed in many recent fluency studies (Housen, Kuiken and Vedder 2012). This means that we distinguish between two main dimensions of fluency: speed and breakdown fluency. Speed fluency is measured on the basis of articulation rate

(AR) which reflects the number of syllables divided by speaking time. Following on from the observations of Goldman-Eisler (1968), De Jong et al (2012: 124) argue that articulation rate constitutes a reliable fluency measure unaffected by situational and individual variability. However, speech rate of samples, which is expressed in the number of syllables per time unit and combines aspects of both speed and breakdown fluency was also established as it is regarded as one of the most reliable predictors of fluency (Kormos 2006: 162). Breakdown fluency refers to the frequency and length of pausing as expressed in measures such as: mean length of runs (MLR), mean length of pauses (MLP) and phonation time ratio (PhTR) denoting the total length of speech divided by the total time spent speaking) (Kormos 2006: 163). MLR is a measure of pause frequency in relation to words produced and is established by dividing the total of syllables by the total of runs. To arrive at a complete set of temporal measurements, a range of temporal input values was initially established for each sample including: overall speech time, pausing time, total of pauses (filled and silent) and total of runs between pauses. This was followed by computing the following set of fluency measures: AR, SR, MLR, MLP and PhTR.

### 5.2.2 Identification of formulaic sequences in the data

The second part of the procedure involved the identification of formulaic sequences in the collected data. The major problem inherent in extracting formulaic sequences from text in a consistent way is that their only truly essential characteristic (that could serve as the ultimate proof of the formulaic status) i.e. their holistic handling, is very difficult, if not impossible, to determine. In the current state of research there is no direct, unambiguous method of ascertaining that a string is stored and handled as a whole (Wray 2008: 113). Taking into account that formulaic status is not always signalled in a straightforward, objectively measurable way, it is particularly difficult to distinguish formulaic strings from grammatically regular and semantically transparent novel sequences. One way to overcome this procedural impasse is to adopt an eclectic approach (Read and Nation 2004: 33). In the absence of a single unquestionable indicator of formulaicity, the researchers have no alternative but to make use of the available criteria derived on the basis of many years of observation and research. Applying mixed criteria that originate from various areas of research and more than one mode of analysis has the advantage of increasing the likelihood of fishing out the majority of formulaic sequences from the dataset.

Keeping the above observations in mind, for the purpose of this study we have employed an identification procedure developed by Erman and Warren (2000), which relies on restricted exchangeability (also referred to as non-substitutability) of the constituent elements of a sequence as the main determinant of formulaic status (see also Wiktorsson 2001, 2003; Forsberg 2006; Forsberg and Fant 2010; Knutsson 2006). Restricted exchangeability expresses the degree of fixedness of a given sequence i.e. the fact that it displays some degree of resistance to change. Applying a restricted exchangeability test involves substituting at least one of the elements of a candidate string with a synonym and analyzing the resulting sequence in terms of changes in meaning, function or idiomaticity. For example, replacing the word 'a doll' in the

expression 'Thanks, you're a doll!' ('You are very nice. Thank you.'), which is used to express gratitude for somebody's help or kindness, with 'a puppet' or 'a girl's toy' results in a radical change of meaning and the loss of the pragmatic function and idiomaticity. Borderline candidate strings (ones for which restricted exchangeability fails to produce consistent results) are subjected to a further verification procedure which involves testing them in terms of the twelve diagnostic formulaicity criteria outlined in Wray and Namba (2003) (see Appendix 2 for a complete list of their criteria). These features can be described as corresponding to the four major characteristics of formulaic language identified in available research i.e. form, meaning, function and provenance (origin) (Wray 2002).

To sum up, our identification procedure for extracting formulaic sequences consists of three stages. Firstly, samples are examined in search of sequences which appear to be formulaic and candidate formulaic strings are selected. The candidate strings are then subjected to restricted exchangeability test. If they are marked positive for restricted exchangeability, they are listed in the inventory of formulaic sequences. If the restricted exchangeability fails or cannot be applied, the candidate strings need to undergo further verification, which involves validating them in terms of Wray and Namba's criteria. Once the formulaic strings are successfully identified, their total in each sample is calculated.

## 5.3 Results and discussion

The central issue addressed in this study concerns the validity of the claim that the use of formulaic language in native speech is linked to a speaker's productive fluency. As already indicated, the apparent advantage of using automatized and/or proceduralized prefabricated, formulaic strings comes from the fact that their holistic retrieval decreases the processing load involved in online speech production. This grants the speaker with additional time and cognitive resources, which can be delegated to handle other aspects of speech (such as discourse planning or social interaction). As a result, speech is produced more smoothly and rapidly. Following this line of reasoning, higher incidence of formulaic sequences in speech should be linked with an increase in the speed of speaking and the length of runs and a decrease in the length and frequency of pausing.

To address the research questions formulated for this study, Pearson product-moment correlation coefficients were computed as a way of investigating the nature and strength of the relationship between the number of formulaic expressions identified in the speakers' output and a range of selected speed and breakdown fluency variables. Table 1 summarises the results.

| n=45 | temporal measure | strength of relationship | significance level |
|---|---|---|---|
| speed fluency | AR | r =.266 | p = n.s |

| n=45 | temporal measure | strength of relationship | significance level |
|---|---|---|---|
| breakdown fluency | MLR | r. = .467 | p ≤ .05 |
| | MLP | r = -. 291 | p = n.s. |
| | PhTR | r =.402 | p ≤ .05 |
| speed and breakdown fluency | SR | r =.457 | p ≤ .05 |

**Table 1**: Formulaic sequences vs productive fluency

The first research question put forward for this study concerned the relationship between speed fluency (as measured by AR) and the use of formulaic sequences in native Polish speech. As illustrated in Table 1, articulation rate was reported not to correlate significantly with the number of formulaic strings found in the sample (r=.266, p=0.38, with the significance level set at .05). As regards speech rate, which combines speaking speed (AR) with overall pause time, a correlation for the data revealed that speech rate and the number of formulaic sequences were significantly related (r=.457). In other words, as predicted in the earlier parts of this paper, increases in the number of formulaic sequences were linked with increases in speech rate scores suggesting that participants whose output abounded in formulaic language were able to speak faster. This provides some preliminary evidence for the fluency-enhancing function of formulaic language with relation to the speed of speech production.

Our second research question was aimed at investigating the relationship between the use of formulaic sequences and the speaker's breakdown fluency (as measured by MLR, MLP and PhTR). Mean length of fluent runs (MLR) was found to correlate significantly with the number of formulaic strings (r=.467). This correlation indicates that a higher incidence of formulaic strings in speech was linked to an increase in the length of pause-free stretches. For mean length of pause, the correlation returned non-significant values (r=-.291, p=0.71). While the correlation was not significant relative to the standard alpha level of .05, the p-value was less than .10 suggesting that this weak negative correlation approaches significance. The existence of a negative correlation between the use of formulaic language and MLP would suggest that a higher number of prefabs in output is associated with a decrease in the mean length of pause. Combined with the correlation values for MLR, these findings could suggest that speakers relying on a more extensive use of formulaic strings produce discourse marked by fewer and shorter pauses.

However, as already stated, the MLP results need to be approached with caution as this correlation returned marginally significant values. The final measure employed as a gauge of breakdown fluency is that of phonation time ratio which, as De Jong et al. (2012: 124) point out, is a single measure which subsumes all the measures of silent pausing - MLR, MLP and the number of pauses. Overall, our results indicate a moderate positive correlation (r=.402) between PhTR and the number of uttered formulaic sequences. To put it differently, increases in the number of formulas used are linked to a higher proportion of time filled with speech.

Taken together, the correlation data obtained in the course of our analysis provide some useful insights into the issue of fluency-enhancing function of formulaic language. As regards both of the investigated dimensions of fluency – speed and breakdown fluency – we found that the use of formulaic strings in the oral native production of Polish native speakers is positively and significantly correlated with mean length of runs, speech rate and phonation time ratio. As we have seen, the data are consistent across the three fluency variables, all of which returned significant and positive values. It is worth pointing out that these three measures (MLR, PhTR and SR) capture the major temporal aspects of speech such as time spent speaking (as opposed to pausing) and the frequency and duration of pauses. Also, it is SR and MLR which are considered the most consistent and reliable temporal indicators of fluency (Kormos and Denes 2004; Wood 2004; Kormos 2006; Préfontaine 2010). On the basis of the speed and breakdown fluency data, it can be concluded that speech characterised by a higher incidence of prefabricated language consists of longer uninterrupted stretches of discourse and is produced with a higher number of words/syllables uttered per time unit. In short, speech which was rich in formulaic language is produced more rapidly and with fewer hesitation phenomena.

However, it has to be noted that although our results provide some preliminary evidence validating the link between formulaicity and fluency in L1 oral performance, at the same time we have not managed to gather a convincing body of data that would substantiate the claims concerning the relatively high strength of this relationship. In fact, the established correlation values did not exceed r=.5, which indicates a relatively weak relationship between the investigated variables. Additionally, two out of the five investigated fluency measures returned non-significant values. Although the number of participants was higher than in most previous research, the limitation of this study is its small sample size. However, the major issue facing any analysis of pause distribution in a larger body of naturally occurring discourse (native or non-native) is the lack of reliable pause annotation in the available spoken corpora (Dahlmann, Adolphs and Rodden 2007). The current annotation conventions fail to accommodate a number of essential aspects of pausing such as: the distinction between silent and filled pauses and pause duration (which are relevant in fluency research). Finally, some objections might be raised in connection with the identification procedure used to extract formulaic sequences from text, which involves some degree of subjectivity.

In sum, it can be stated that our research has lent some support to the claims concerning the fluency-enhancing function of formulaic language. From a psycholinguistic perspective, formulaic sequences provide adult language users with a kind of 'a shortcutting device' (Peters 1983: 2) by reducing the processing load involved in the production of speech. The considerable computational power involved in constructing each utterance afresh leads native speakers to "economize by stitching

together language chunks (…) so that planning for the form and content of future utterances proceeds more smoothly" (Skehan 1998: 4). Relying on formulaic language allows language users to save planning time and processing effort and direct their attention to other aspects of speech. The major contribution of this study is that it relates the use of formulaic language to a set of objectively measured temporal fluency measures and so, provides a clear, numerical illustration of the speaker's processing gains.

In fact, ample evidence for the processing advantage of formulaic language comes from research on highly conventionalized speech of so called 'smooth talkers' (Schmitt and Carter 2004: 5) whose performance involves the need to maintain fluency in the face of considerable overload of working memory and additional external pressures. Kuiper (2004: 43-45) provides an extensive overview of relevant work in this area, which includes analyses of the performance of auctioneers, sport commentators of fast-paced sports such as ice-hockey and horse-racing and slow-paced sports such as cricket, hockey, weather forecasters and check-out assistants. The professionals from within all of these occupations are required to memorize large stocks of highly detailed information and engage in a kind of non-stop on-line processing to maintain a flow of fluent speech (be it a commentary, a forecast or a sales talk) in the face of additional pressures such as unexpected events, temporal constraints, the response of the audience etc. Kuiper's (2004) analysis of the speech of such live performers reveals that they avoid using novel formulations and tend to rely on formulaic speech which allows them to convey the maximum of information in the minimum of time, at the same time enabling them to maintain pause-free, non-hesitant speech.

## 6. Conclusion

The study reported in this paper investigated the relationship between L1 formulaicity and fluency. In the research we analysed speech samples collected from 45 native speakers of Polish in terms of the formulaic language used and temporal fluency. Pearson product-moment correlations were established to investigate the nature and strength of the link between the participants productive fluency and their use of formulaic sequences. The use of formulaic language was found to be positively and significantly correlated with speech rate, mean length of runs and phonation time ratio. This suggests that a higher concentration of formulaic material in participants' output is associated with faster speed of speech, longer stretches of speech between pauses and an increase in the amount of time filled with speech. No statistically significant correlations were found between participants' use of formulaic language and the articulation rate and mean length of pauses. To conclude, in our view this investigation provides some preliminary evidence for the facilitating role of formulaic sequences with relation to the speed of speech production and the occurrence of hesitation phenomena. However, it is recommended that the study is replicated with a larger body of data to provide further insight into the issue.

# References

Aijmer, K. 1996. Conversational Routines in English. Longman: London and New York.

Altenberg, B. 1990. Speech as linear composition. In Caie, G. Haastrup, K., Jakobsen, A.L., Nielsen, J.E., Sevaldsen, J., Specht, H. and A. Zettersten, (eds.) Proceedings from the Fourth Nordic Conference for English Studies, Vol. 1. Department of English, University of Copenhagen: 133–143.

Anderson. J. R. 1983. The Architecture of Cognition. Cambridge, Massachusetts: Harvard University Press.

Biber, D., S. Johansson, S. G. Leech, S. Conrad, and E. Finegan. 1999. Longman Grammar of Spoken and Written English. London: Longman.

Boersma, P. and D. Weenink. 2005. PRAAT. Retrieved June, 2013 from http://www.praat.org.

Chambers, F. 1997. What do we mean by fluency? System 25(4): 535–544.

Corrigan, R., E. A. Moravcsik, H. Ouali, and K. M. Wheatley. 2009a. Formulaic Language: Volume 1: Distribution and Historical change. Amsterdam: John Benjamins.

Corrigan, R., E. A. Moravcsik, H. Ouali, and K. M. Wheatley. 2009b. Formulaic Language: Volume 2: Acquisition, Loss, Psychological Reality, and Functional Explanations. Amsterdam: John Benjamins.

Cowie, A. P. 1998. Phraseology: Theory, Analysis and Applications. Oxford: Clarendon Press.

Dahlmann, I., Adolphs, S. and T. Rodden. 2007. Multi-word expressions fluency and pause annotation in spoken corpora. Paper presented 40th BAAL Annual Meeting on Technology, Ideology and Practice in Applied Linguistics, Edinburgh, UK, September 6–8, 2007.

De Jong, N., Halderman, L.K. and M. Ross. 2009. The effect of formulaic sequences training on fluency development in an ESL classroom. Paper presented at the American Association for Applied Linguistics conference 2009, Denver, CO, March 2009.

De Jong, N. H., M. P. Steinel, A Florijn, R. Schoonen, and J. H. Hulstijn. 2012. The effect of task complexity on functional adequacy, fluency and lexical diversity in speaking performances of native and non-native speakers. In A. Housen, F. Kuiken and I. Vedder (eds.) Dimensions of L2 Performance and Proficiency: Complexity, Accuracy and Fluency in SLA. Amsterdam: John Benjamins:121–142.

Dechert, H. W. 1984. Second language production: six hypothesis. In H. W. Dechert, D. Möhle and M. Rapuach (eds.) Second Language Productions. 211–230. Tübingen, Germany: Gunter Narr Verlag.

Eeg-Olofsson, M. and B. Altenberg. 1996. Recurrent word combinations in the London Lund Corpus: coverage and use of word-class tagging." In C. E. Percy, C. E. Meyer, and I. Lancashire (eds.) Synchronic Corpus Linguistics: Papers from the Sixteenth International Conference on English Language Research on Computerized Corpora (ICAME 16). Amsterdam: Rodopi B.V.: 97–108.

Erman, B. and B. Warren. 2000. The idiom principle and the open choice principle. Text, 20 (1): 29–62.

Fillmore, C. J. 1979. On fluency. In D. Kempler and W. S. Y. Wang (eds.) Individual differences in language ability and language behavior, New York: Academic Press: 85–102.

Forsberg, F. 2006. Le Langage Préfabriqué en Francais Parlé L2. PhD Thesis, Stockholm University, Stockholm.

Forsberg, F., and L. Fant. 2010. Idiomatically speaking – effects of task variation on formulaic language in high proficient users of L2 French and Spanish. In D. Wood (ed.), Perspectives on Formulaic Language in Acquisition and Communication, New York: Continuum: 47–70.

Freed, B. 2000. Is fluency in the eyes (and ears) of the beholder? In H. Riggenbach (ed.) Perspectives on Fluency. University of Michigan Press: Ann Arbor: 243–265.

Freed, B. F., N. Segalowitz, and D. P. Dewey. 2004. Context of learning and second language fluency in French: comparing regular classroom, study abroad, and intensive domestic immersion programs. Studies in Second Language Acquisition 26(02): 275–301.

Gatbonton, E. and Segalowitz, N. 1988. Creative automatization: principles for promoting fluency within a communicative framework. TESOL Quarterly, 22 (3): 473–492.

Gatbonton, E. and Segalowitz, N. 2005. Rethinking communicative language teaching: a focus on access to fluency. Canadian Modern Language Review, 61: 325–353.

Goldman-Eisler, F. 1968. Psycholinguistics: Experiments in Spontaneous Speech. London: Academic Press.

Götz, S. 2013. Fluency in Native and Nonnative English Speech. Amsterdam: John Benjamins.

Guillot, M.-N. 1999. Fluency and its Teaching. Clevedon, UK: Multilingual Matters.

Housen, A., Kuiken, F. and I. Vedder. 2012. Dimensions of L2 Performance and Proficiency: Complexity, Accuracy and Fluency in SLA. Amsterdam/Philadelphia: John Benjamins Publishing.

Hunston, S. and Francis, G. 2000. Pattern Grammar: A Corpus-Driven Approach to the Lexical Grammar of English. Amsterdam: John Benjamins.

Knutsson, R. 2006. Formulaic Language in L1 and L2. Licentiate Dissertation. Lund University, Lund.

Kormos, J. 2006. Speech Production and Second Language Acquisition. Mahwah, N.J.: Lawrence Erlbaum.

Kormos, J. and Dénes, M. 2004. Exploring measures and perceptions of fluency in the speech of second language learners. System, 32: 146–164.

Kuiper, K. 2004. Formulaic performance in conventionalized varieties of speech. In N. Schmitt (ed.) Formulaic Sequences: Acquisition, Processing and Use. Amsterdam: Benjamins: 37–54.

Lennon, P. 2000. The lexical element in spoken second language fluency. In H. Riggenbach (ed.). Perspectives on Fluency. Ann Arbor: The University of Michigan Press: 25-42.

Levelt, W. 1989. Speaking from Intention to Articulation. Cambridge, MA: The MIT Press.

Meunier, F. and S. Granger. 2008a. Phraseology in Foreign Language Learning and Teaching. Amsterdam: John Benjamins.

Meunier, F. and S. Granger. 2008b. Phraseology: an Interdisciplinary Perspective. Amsterdam: John Benjamins.

Moon, R. 1998. Fixed Expressions and Idioms in English. A Corpus Based Approach. Oxford Studies in Lexicography and Lexicology. Oxford: Oxford University Press.

Nattinger, J. and J. DeCarrico. 1992. Lexical Phrases and Language Teaching. Oxford: Oxford University Press.

Pawley, A. 2009. Grammarians' languages versus humanists' languages and the place of speech act formulas in models of linguistic competence. In R. Corrigan, E. A. Moravcsik, H. Ouali, and K. M. Wheatley (eds) Formulaic Language: Volume 1: Distribution and Historical Change. Amsterdam: John Benjamins: 3–26.

Pawley, A. and F. H. Syder. 1983. Two puzzles for linguistic theory: nativelike selection and nativelike fluency. In J. C. Richards, and R. W. Schmidt (eds.) Language and Communication. London: Longman: 191–225.

Pawley, A. and F. H. Syder. 2000. The one-clause-at-a-time hypothesis. In H. Riggenbach (ed.), Perspectives on fluency. Ann Arbor: University of Michigan Press: 163–199.

Peters, A. M. 1983. The Units of Language Acquisition. Cambridge: Cambridge University Press.

Préfontaine, Y. 2010. Differences in Perceived Fluency and Utterance Fluency across Speech Elicitation Tasks: A Pilot Study. Papers from Lancaster Univesrity Postgraduate Conference in Linguistics and language Teaching, 5:134–154.

Raupach, M. 1980. Temporal variables in first and second language speech production. In H.W. Dechert and M. Raupach (eds.) Temporal Variables in Speech. The Hague: Mouton: 263–270.

Read, J. and Nation, P. 2004. Measurement of formulaic sequences. In N. Schmitt (ed.) Formulaic Sequences: Acquisition, Processing and Use. Amsterdam: John Benjamins Publishing, 23-35.

Renouf, A. and J. Sinclair. 1991. Collocational frameworks in English. In K. Aijmer and B. Altenberg (eds.) English Corpus Linguistics: Studies in the Honour of Jan Svartvik. London, Longman: 128–143.

Schmitt, N. 2004. Formulaic Sequences. Amsterdam: John Benjamins.

Schmitt, N. and R. Carter. 2004. Formulaic sequences in action: an introduction. In N. Schmitt (ed.) Formulaic Sequences: Acquisition, Processing and Use. Amsterdam: John Benjamins: 1–22.

Segalowitz, N. 2003. Automacity and second languages. In C. Doughty and M. Long (eds.) The Handbook of Second Language Acquisition. Oxford, UK: Blackwell: 382–408.

Segalowitz, N. 2010. The Cognitive Bases of Second Language Fluency. New York: Routledge.

Segalowitz, N. and B. F. Freed. 2004. Context, contact and cognition in oral fluency acquisition: learning Spanish in at home and study abroad contexts. Studies in Second Language Acquisition 26(2): 173–199.

Sinclair J. 1991. Corpus, Concordance, Collocation. Oxford: Oxford University Press.

Skehan, P. 1998. A Cognitive Approach to Language Learning. Oxford: Oxford University Press.

Skehan, P. 2003. Task based instruction. Language Teaching, 36 (1): 1–14.

Skehan, P. 2009. Modelling second language performance: integrating complexity, accuracy, fluency, and lexis. Applied Linguistics 30(4): 510–532.

Tavakoli, P. and P. Skehan. 2005. Strategic planning, task structure, and performance testing. In R. Ellis (ed.) Planning and Task Performance in a Second Language. Amsterdam: John Benjamin: 239–273

Towell, R., Hawkins, R. and N. Bazergui. 1996. The development of fluency in advanced learners of French. Applied Linguistics 17: 84–119.

Weinert, R. 2010. Formulaicity and usage-based language: linguistic, psycholinguistic and acquisitional manifestations. In D. Wood (ed.) Perspectives on Formulaic Language. Acquisition and Communication. London/New York: Continuum: 1–22.

Wiktorsson, M. 2001. Register Differences between Prefabs in Native and EFL English. The Department of English in Lund: Working Papers in Linguistics 1: 85–94. Retrieved January 2009 from
http://www.sol.lu.se/engelska/dokument/wp/vol01/Maria.pdf .

Wiktorsson, M. 2003. Learning Idiomaticity: A Corpus-based Study of Idiomatic Expressions in Learners' Written Production. Lund Studies in English. Vol. 105. Lund: Lund University.

Wood, D. 2001. In search of fluency: What is it and how can we teach it? Canadian Modern Language Review, 57: 573–589.

Wood, D. 2004. An empirical investigation into the facilitating role of automatized lexical phrases in second language fluency development. Journal of Language Learning, 2(1): 27–52.

Wood, D. 2006. Uses and functions of formulaic sequences in second language speech: an exploration of the foundations of fluency. The Canadian Modern Language Review 63: 13–33.

Wood, D. 2007. Mastering the English formula: fluency development of Japanese learners in a study abroad context. JALT Journal 29 (2): 209–230.

Wood, D. 2009. Effects of focused instruction of formulaic sequences on fluent expression in second language narratives: A case study. Canadian Journal of Applied Linguistics, 12 (1), 39–57.

Wood, D. 2010a. Perspectives on Formulaic Language. Acquisition and Communication. London/New York: Continuum.

Wood, D. 2010b. Formulaic Language and Second Language Speech Fluency: Background, Evidence, and Classroom Applications. London/New York: Continuum.

Wray, A. 2002. Formulaic Language and the Lexicon. Cambridge: Cambridge University Press.

Wray, A. 2008. Formulaic Language: Pushing the Boundaries. Oxford: Oxford University Press.

Wray, A. and K. Namba. 2003. Formulaic language in a Japanese-English bilingual child: a practical approach to data analysis. Japan Journal for Multilingualism and Multiculturalism 9 (I): 24–51.

Wray, A. and M. R. Perkins 2000. The functions of formulaic language: an integrated model. Language and Communication 20(1): 1–8.

# Appendix 1

1. Opowiedz o najbardziej cennej rzeczy jaką posiadasz lub przedmiocie, do którego jesteś szczególnie przywiązana/y. Adapted from Inside Out, Student's Book Upper intermediate p. 32, own translation.
2. Opowiedz o tym jak wyglądało twoje życie w wieku około 8 lat. Adapted from Inside Out, Student's Book Intermediate p. 49, own translation.
3. Opowiedz o sytuacji, kiedy robiła/eś coś po raz pierwszy. Adapted from Inside Out, Student's Book Advanced p. 87, own translation.
4. Opowiedz o momencie, w którym poczuła/eś przepływ adrenaliny. Adapted from Inside Out, Student's Book Intermediate p. 38, own translation.
5. Opowiedz o filmie, który zrobił na Tobie duże wrażenie lub kompletnie cię rozczarował. Adapted from Inside Out, Student's Book Upper intermediate p. 108, own translation.

# Appendix 2

The Eleven Diagnostic Criteria for Identification of Formulaic Sequences. Adapted from Wray and Namba (2003).

1. A: By my judgement there is something grammatically unusual about this wordstring.
2. B: By my judgement, part or all of the wordstring lacks semantic transparency.
3. C: By my judgement, this wordstring is associated with a specific situation and/or register.
4. D: By my judgement, the wordstring as a whole performs a function in communication or discourse other than, or in addition to, conveying the meaning of the words themselves.
5. E: By my judgement, this precise formulation is the one most commonly used by this speaker/writer when conveying this idea.
6. F: By my judgement, the speaker/writer has accompanied this wordstring with an action, use of punctuation, or phonological pattern that gives it special status as a unit, and/or is repeating something s/he has just heard or read.
7. G: By my judgement, the speaker/writer, or someone else has marked this wordstring grammatically or lexically in a way that gives it special status as a unit.
8. H: By my judgement, based on direct evidence or my intuition, there is a greater than-chance-level probability that the speaker/writer will have encountered this precise formulation before, from other people.
9. I: By my judgement, although this wordstring is novel, it is a clear derivation, deliberate or otherwise, of something that can be demonstrated to be formulaic in its own right.
10. J: By my judgement, this wordstring is formulaic, but it has been unintentionally applied inappropriately.

11. K: By my judgement, this wordstring contains linguistic material that is too sophisticated, or not sophisticated enough, to match the speaker's general grammatical and lexical competence.