Cambridge English Centenary
Symposium on Speaking Assessment

Fluency

Glenn Fulcher, University of Leicester

Fluency has always been challenging to define and operationalize in speaking tests. Yet, it persists as a concept with which teachers are comfortable, and a construct that language testers find irresistible. Brumfit (1984) characterised fluency as existing in a relationship of polarity with 'accuracy', describing it as "natural language use" (1984:56). Drawing on earlier work by Fillmore (1979), he characterized fluency as (a) "filling time with talk", which implies automaticity of language processing; (b) the production of coherent sentences using the "semantic and syntactic resources of the language" appropriately; (c) selecting appropriate content for context; and (d) being creative with the language. When everything comes together with appropriate world knowledge and suitable personality, a speaker is rarely "lost for words" and does not become "tongue tied". We get the impression that speech "flows". Koponen and Riggenbach (2000) deconstruct the metaphorical understanding of fluency as "language as motion"; fluid like liquid, or flowing like a river. Speech is said to be "smooth, rapid and effortless", rather than "choppy" (Chambers 1997).

This metaphor reveals that fluency is situated partially in the production of the speaker, and partially in the perception of the listener (Freed 2000). In language testing the construct is articulated in rating scale descriptors, critiques of which became common in the 1980s (Lantolf and Frawley 1985, 1988). Based primarily on principles of internal coherence of the "more than/less than" type, qualifiers such as "*undue* hesitation" and "*excessive* pausing" also invite listener comparison with some internalised abstraction of an ideal "native speaker" (Davies 2004). In an attempt to overcome these problems Fulcher (1987, 1993, 1996) advocated the development of data-based scales to generate descriptors that were grounded in learner performance. The top and bottom descriptors were now defined by the most and least fluent performances on test tasks, and descriptors were generated from the analysis of observable performance features.

Fluency research has since broadly taken two paths: the cognitive science route and the linguistic route. Both aspire to describe the observable features of fluency in speech. After all, fluency is a construct, and must have observable elements that together define that construct for it to be measurable. Both wish to understand what raters pay attention to when making judgments about fluency. However, they differ in one critical respect. Cognitive science researchers add "L2 cognitive fluency" to the mix (Segalowitz 2010:76). It is said to be the cause of the observable features of (dys)fluency in performance. Researchers therefore wish to discover (a) which features most impact on the perceived fluency of a listener, and (b) which features can be predicted by variation in cognitive fluency. Each observable feature is therefore treated as a variable capable of objective measurement that *does not in itself need interpretation*, but is explained in terms of the effects of another construct. The practical effect is that counts of observations may be fed directly into regression models following removal of outliers and data normalization.

The linguistic school accepts the relevance of cognitive speech models for understanding language processing and production (Levelt 1989, 1999, Field 2011). However, the explanation for all surface phenomena is not necessarily cognitive. The cognitive school pays scant attention to language as a means of communication that is largely conditioned by social interaction. From the linguistic perspective it is argued that there is no single explanation for any dysfluency feature or change in speech rate. Nor is there one-to-one mapping with components of an L2 cognitive fluency model. These features can be measured, but in some contexts they will be perceived as dysfluency, and in others as quite fluent. Speakers deliberately use silence to communicate their impressions, attitudes, emotions, and intentions (Nakane 2007, Bruneau 2008), and listeners interpret speed and pauses in these terms. Pauses are also part of our turn-management toolkit (McCarthy, 2010), and a politeness mechanism (Scollon and Scollon 1989). They are a vehicle for expressing our personality, establishing social status, and injecting suspense or comic relief into utterances (Nakane 2012). The argument from the linguistic school is that surface features must be interpreted in context, as their manifestation will vary according to task and task features (Skehan 2009) as well as speaker intention.

The choice of approach also has profound implications for how we assess fluency. If one holds that cognitive fluency can "… serve as a stand-in measure of general proficiency and L2 experience" (Segalowitz 2010:76), no interpretation of observable features is necessary. They become the indirect measure of L2 proficiency by virtue of their relation with cognitive fluency, and can be measured by current computer technology (De Jong and Wempe 2009, Little et al 2013). Tasks like sentence repetition, read-aloud, and sentence building (Van Moere 2012) provide the data necessary for the psycholinguistic/cognitive inferences that are said to allow extrapolation to claims about likely candidate performance beyond the test. Alternatively, if language performance is more than a cognitive process that involves interactive responsiveness and the use of a rich repertoire of linguistic choices to express meaning, the use of human judgment seems to be an inevitable requirement.

In this talk I consider the range of fluency/dysfluency phenomena that have been studied, and summarize key findings. I discuss whether the cognitive or linguistic path provides the most convincing interpretation of data in the light of models of speech processing. I then look at how fluency has been articulated in rating scales to show that these have been more successful than is commonly thought to be the case. I conclude with a rationale for persisting with the use of human ratings in assessing spoken fluency.

References

Brumfit, C (1984) *Communicative methodology in language teaching: The roles of fluency and accuracy*, Cambridge: Cambridge University Press.

Bruneau, T J (2008) How Americans use silence and silences to communicate, *China Media Research*, 4(2), 77-85.

Chambers, F (1997) What do we mean by fluency? *System 24*(4), 535-544.

Davies, A (2004) The native speaker in applied linguistics, in Davies, A and Elder, C (Eds.), *The Handbook of Applied Linguistics*, Oxford: Blackwell, 431-450.

De Jong, N and Wempe, T (2009) Praat script to detect syllable nuclei and measure speech rate automatically, *Behavior Research Methods,* 41,385-390.

Field, J (2011) Cognitive Validity, In Taylor, L. (Ed.) *Examining speaking: Research and practice in assessing second language speaking,* Cambridge: UCLES/Cambridge University Press, 65-111.

Fillmore, L W (1979) Individual differences in second language acquisition, in Fillmore, C, Kempler, D and Wang, W Y S (Eds) Individual Differences in Language Ability and Language Behavior, New York: Academic Press.

Freed, B F (2000) Is fluency, like beauty, in the eyes (and ears) of the beholder? In Riggenbach, H (Ed.), *Perspectives on fluency*, Ann Arbor: University of Michigan Press, 243-265.

Fulcher, G (1987) Test of oral performance: The need for data-based criteria, *English Language Teaching Journal*, 41(4), 287-291.

Fulcher, G (1993) The Construction and Validation of Rating Scales for Oral Tests in English as a Foreign Language, unpublished PhD thesis, University of Lancaster, UK.

Fulcher (1996) Does thick description lead to smart tests? A rating- scale approach to language test construction, Language Testing 13 (2), 208–38.

Koponen, M and Riggenbach, H (2000) Overview: Varying perspectives on fluency, in Riggenbach, H (Ed.), *Perspectives on fluency,* Michigan: University of Michigan Press, 5-24.

Lantolf, J P and Frawley, W (1985) Oral proficiency testing: A critical analysis, *The Modern Language Journal*, 69, 337-345.

Lantolf, J P and Frawley, W (1988) Proficiency: Understanding the construct, *Studies in Second Language Acquisition*,10, 181-196.

Levelt, W J M (1989) Speaking, Cambridge, MA: MIT Press.

Levelt, W J M (1999) Language production: a blueprint of the speaker, in Brown, C and Hagoort, P (Eds) Neurocognition of Language, Oxford: Oxford University Press, 83–122.

Little, D R, Oehmen, R, Dunn, J, Hird, K and Kirsner, K (2013) Fluency profiling system: An automated system for analysing the temporal properties of speech, *Behavior Research Methods* 45(1), 191 – 202.

McCarthy, M (2010) Spoken fluency revisited, *English Profile Journal*, 1, e4.

Nakane, I (2007) *Silence in intercultural communication: Perceptions and performance,* Amsterdam: John Benjamins.

Nakane, I (2012) Silence, in Paulston, C B, Kiesling, S F, and Rangel, E S (Eds.), *Handbook of intercultural discourse and communication*, Oxford: Blackwell, 158-179.

Scollon, R and Scollon, S (1989) *Narrative, literacy and face in interethnic communication*. Norwood, New Jersey: Ablex.

Segalowitz, N (2010) *The cognitive bases of second language fluency*, New York: Routledge.

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

Van Moere, A (2012) A psycholinguistic approach to oral language assessment, *Language Testing*, 29(3), 325-344.