ROBIN LICKLEY

Disfluency in typical and stuttered speech

This paper discusses what happens when things go wrong in the planning and execution of running speech, comparing disfluency in typical speech with pathological disfluency in stuttering. Spontaneous speech by typical speakers is rarely completely fluent. There are several reasons why fluency can break down in typical speech. Various studies suggest that we produce disfluencies at a rate of around 6 per 100 fluent words, so a significant proportion of our utterances are disfluent in some way. Stuttering can halt the flow of speech at a much higher rate than typical disfluency. While persons who stutter are also prone to the same kinds of disfluency as typical speakers, their impairment results in the production of other forms of disfluency that are both quantitatively and qualitatively different from typical forms. In this paper, I give an overview of the causes of disfluency in both typical and stuttered speech and relate these causes to their articulatory and phonetic realisations. I show how typical and stuttered disfluencies differ in both their cause and their realisations.

Key words: fluency, disfluency, stuttering.

1. *Introduction*

This work discusses phenomena that occur in everyday speech, both for typical speakers and for speakers who stutter: Disfluencies. We begin with an operational definition of fluency, before examining how planning difficulties can lead to a failure to maintain a perfectly smooth flow. Types and causes of typical disfluencies are described before we move on to consider disfluencies that occur in stuttered speech.

2. *Fluency*

Before we begin to discuss disfluency in speech, it is important to define what we mean by fluency. Essentially, the word fluency refers to the flow of speech. Flow involves regular, continuous motion over time. If speech is fluent, then the sounds that we hear are perceived as flowing smoothly and without unexpected breaks.

Some breaks in the flow are to be expected in extended periods of speech and these do not necessarily appear disfluent. Speakers need time to breathe in from time to time, after all.

Breath pauses tend to fall at structural boundaries. So, in fluent speech, pauses in the flow are expected at sentence and phrase boundaries. The pattern holds more for read speech than for spontaneous speech (Wang, Green, Nip, Kent & Kent, 2010) and in faster speech, breath pauses become less regularly associated with structural units (Grosjean, Collins, 1979). But within phrases, words typically

follow on from each other continuously. Sounds in the speech stream influence each other and coarticulate, rather than simply concatenating like typed letters. Fluency between sentences and phrases may entail pauses between structural units, but fluency within phrases involves continuity between words, with gestural overlap between adjacent phonemes.

The production of speech demands a complex interaction of processes on many levels. Fluency at the level of speech output depends on smooth functioning within and between all of those levels, from the conceptual planning of what message to convey and how to express it, through formulation, involving lexical selection and syntax building and phonological encoding, all the way to the creation of a motor plan to suit the message and its articulation via the muscle systems that allow us to speak. Thus, it is possible to consider fluency at multiple levels of production prior to the spoken output. If conceptual planning is straightforward, and the speaker can quickly decide what in general terms to say and how to make it appropriate given the current discourse, then there is no need for hesitation before formulation begins, and conceptualisation can be said to be fluent. At the formulation stage, if the lemmas required to express the planned utterance are easily accessed, if construction of the required syntactic structure, the assignment of appropriate morphology and the phonological encoding and prosodic patterns present no problems, then a prearticulatory plan can be constructed fluently. If motor commands to the multiple muscle systems involved in articulation can be synchronised precisely, without a hitch, then a fluent utterance can be produced.

In reading out loud, and in carefully rehearsed performances, the speaker avoids the need for much of the rapid planning behind the act of speaking, and it is relatively rare for fluency to break down. In spontaneous speech, though, the complexity of the task means that it is rare for a speaker to continue absolutely fluently for long.

3. *Typical Disfluency*

If there is a breakdown in fluency, then we can say that the resultant speech is disfluent. Disfluencies occur frequently in typical spontaneous speech, at a rate of around 6 per 100 words (Bortfeld, Leon, Bloom, Schober & Brennan, 2001; Eklund, 2004; Fox Tree, 1995; Shriberg, 1994). They occur at a higher rate in longer utterances (Oviatt, 1995; Shriberg, 1994) and in more complex utterances (Lickley, 2001; Shriberg, 1994). Individuals vary considerably in the rate at which they produce disfluencies, but it is difficult to find a speaker who is never disfluent.

The word 'disfluency' is defined in several different ways in the research literature and there seems to be no consensus on what phenomena it includes, so it is important to begin this piece with our own operational definition. Our definition of fluency refers to the flow of speech, so disfluency involves a break in that flow, when the speaker stops for a moment in a place or for a length of time not predicted by typical fluent production.

We will now describe what phenomena fall under this definition, with the help of some examples, before discussing how these disfluencies result from problems at various points in the production processes.

From a formal point of view, the major subtypes of disfluency are (1) those that involve simply a halt in the production process, (2) those that involve repetition of part of the utterance and (3) those that involve alteration of part of the utterance. We can refer to these as hesitations, repetitions and repairs.

The simplest forms of disfluencies are hesitations. Hesitations may be realized as silent pauses, prolongations, filled pauses and repetitions, as well as combinations of these. The simplest form of hesitation is a silent pause. Though it may be the simplest form, it is also one of the hardest to define. One problem is that it is difficult to define a minimum duration for what should constitute a disfluent silence. Another is that silences occur in fluent speech, between turns in dialogue (though speakers often overlap), between sentences and phrases (though by no means all of them), and within words, where a short silence can occur during the closure phase of a stop consonant. Distinguishing between fluent and disfluent pauses is far from easy.

On the duration issue, many researchers have accepted 250 ms as a minimum for within-sentence hesitation pauses, following Goldman-Eisler's assertion that only pauses longer than this should be included in accounts of the cognitive processes involved in hesitation (Goldman-Eisler, 1958a; 1958b; 1961). In Goldman-Eisler's view, shorter silences were more likely a consequence of typical fluent production processes. This rather arbitrary value possibly stemmed from the limitations of the technology available at the time for analyzing speech. Since rate of speech varies not only between speakers on their average rate, but also within speakers, depending on discourse context and within single utterances, and in slower speech one would expect longer silences, it should be clear that it is not even desirable to set a single threshold for a hesitation silence. Since Goldman-Eisler, of course, the topic has been addressed by many authors. Among these, Butcher (1981) examined the effect of prosodic context on the perception of pauses in German. Listeners regarded a 220 ms pause as excessive when it occurred between tone groups; but a pause of only 80 ms was regarded as excessive when it occurred within a tone group. How long a pause has to be to be judged as disfluent depends on context, so an absolute value for a hesitation is hard to justify.

On distinguishing fluent from disfluent pauses, we need to consider structural issues at the syntactic and prosodic levels. There is not a straightforward relationship between prosodic and syntactic structure and Ferreira (1993; 2007) suggests that the durations of fluent pauses are licensed by prosodic structure, while hesitation pauses, most often caused by planning issues, are more closely related to syntactic structure, since it is at points related to structural planning and lexical access that speakers need to find time.

In classifying disfluencies within speech corpora, given that there are problems both in defining pause durations and in distinguishing fluent from disfluent pauses, it seems safest, where possible, to employ subjective judgement, supported by

inter-judge agreement, in classifying pauses in running speech, a system used, for example, by Nakatani and Hirschberg (1994) and Eklund (2004).

Aside from these issues with defining silent pause, as intimated previously, it is unclear whether we should distinguish silent from breath pauses. A major issue here is that in analyzing recordings of speech it may not always be possible to judge accurately when a pause includes a breath. But this is a reflection of life beyond the speech laboratory, where such acoustic detail may easily be drowned out by background noise.

Speakers also use prolongation of syllables in otherwise fluent speech as a means of pausing. For the same reasons as those given for silent pauses, prolongations are hard to classify in a rule-based way, since syllables vary in duration in fluent speech according to context. So, again, subjective judgement is often needed. Prolongation may accompany other signs of pausing, preceding a silent pause, for example (Bell, Jurafsky, Fosler-Lussier, Girand, Gregory & Gildea, 2003). On their own, prolonged syllables can give the listener a clear impression that the speaker has halted momentarily, even in the absence of any acoustically measureable silence (Duez, 1993).

By convention, when a speaker is expected to say something, they will say something, if only to let others know that they're about to say something more. A long silence is open to misinterpretation (Does the person realize that it is their turn to speak? Do they have nothing to say?). So a vocal sign of hesitation makes sense from the point of view of continuity within a dialogue. The most recognizable sign of vocalized pausing is what is usually referred to as the filled pause. In English, filled pauses can be represented orthographically as "um" or "uh" (amongst many other variants). The written form disguises the fact that the phonetic realization varies greatly between accents of English. Being syllabic and relatively long, filled pauses are more easily recognised than silent pauses and prolongations, so one can report their frequency in speech corpora with greater confidence. For example, for the 64 speakers in the HCRC Map Task Corpus (Anderson, Bader, Bard, Boyle, Doherty, Garrod, Isard, Kowtko, McAllister, Miller, Sotillo, Thompson & Weinert, 1991, mostly Scottish-accented English) we find an overall average rate of 1.3 filled pauses per 100 words, with individual speakers ranging from 0.18 to 6.66 per 100 words (Lickley, 2015), while reports of other corpora in American English vary between 2.6 (Bortfeld *et al.*, 2001) and 1.6-2.2 (Shriberg, 1994, for 3 different corpora). The differences in frequency between corpora are likely due to task complexity, rather than dialect differences. In languages other than English, different frequencies have been reported. For example, Eklund (2004) reports a rate of 3.6 per 100 words for a Swedish corpus and Maekawa (2004) reports 7.23 per 100 words for a Japanese corpus. It is hard to make valid cross-corpus comparisons and to conclude anything from such differences, when not only the demands of the speaking tasks but also the definitions of 'filled pause' may vary between studies. Japanese filled pauses, for example, include lexical fillers, loosely transcribed as "ano" and "eeto" (Watanabe, 2009), while many studies of 'filled pause' also include prolonged syllables and repetitions under the definition.

It was suggested above that repetitions form a disfluency subtype of their own. This is true in that they are different in form from others, but a case can be made for some repetitions functioning as hesitations and others as repairs. When speakers pause mid-utterance they have the option of continuing from where they left off or restarting the phrase that they paused, thus repeating a word or two. Disfluent repetitions in English tend to be on unstressed function words (Clark, Wasow, 1998; Fox, Jasperson, 1995; Lickley, 1994; Shriberg, 1994). Lickley (1994) found that 96% of disfluent repetitions were on function words (very often the definite article, "the").

When discussing repetitions, we also need to distinguish fluent from disfluent cases. Occasionally, two instances of a word need to be produced together for an utterance to convey the intended meaning: telephone numbers often contain repeated digits, for example. People regularly repeat intensifiers for rhetorical effect (e.g., "that was a very, very strange election result"). Where repetitions form part of a fluent utterance, they will fall into the prosodic pattern expected of the utterance: In saying a telephone number out loud, there is a characteristic intonation pattern, which would be the same whether or not a digit is repeated. Disfluent repetitions are often accompanied by pause of another kind, like silence and prolongation, and they typically repeat the pitch level as well as the word. So, it is relatively easy to distinguish fluent from disfluent cases of repetition, on the basis of prosodic information.

The third major type of disfluency, Repair, can take several forms. Repairs take place when something has gone awry in the production process and the speaker needs to adjust what they have said in order to produce a corrected version of the utterance. This entails a certain amount of backtracking. Most relatively recent corpus-based accounts of disfluency include the repair subtypes Substitution, Insertion and Deletion, as well as combinations of these, and complex cases (Eklund, 2004; Heeman, 1997; Lickley, 1998; Shriberg, 1994).

In Substitutions, an erroneous string is replaced by a correction. The error may be at one of a number of different levels. In this example, the non-word aphosia is produced via a phonological error – anticipation of the stressed vowel from the following word – which is quickly repaired: "Then you can calculate the aphosia – aphasia quotient". In the following example, the speaker selects the wrong lexical item and replaces it after backtracking to the verb: "Go along the road and turn left – turn right at the traffic lights".

In Insertions, the speaker adds something to what they have already said, usually to be more specific. Here, the speaker chooses to add the modifier sharp to make the instruction clearer: "After you reach the post office, take a right – a sharp right turn".

In deletions, the speaker abandons the utterance altogether and immediately begins a new utterance, without apparently altering or adding to the original utterance: "Go along the road and turn – have you visited Edinburgh before?" In this case, the speaker changes plan, deciding to verify that the listener possesses relevant background knowledge before issuing further instructions.

As mentioned above, repetitions can also sometimes viewed as a manifestation of repair. This follows from the view that errors may be detected in the production process prior to being articulated, covertly in other words. If the monitoring and repairing processes are rapid enough, there may be no overt production of the error. But since it takes a certain amount of time to make the correction, the speaker may repeat the onset of the phrase containing the repaired error. Levelt (1983) introduced the notion of covert error and repair, and since then researchers have included repetitions amongst the phenomena that signal covert repair (The Covert Repair Hypothesis: Postma, Kolk, 1993). There is some evidence that there are acoustic differences between types of repetitions, making it possible to distinguish repetitions that are associated with covert repair from others (Plauché, Shriberg, 1999).

Having introduced the major forms of disfluencies that are typically found in spontaneous speech, we now need to discuss in more detail why they occur.

Preparing to speak can take time. As speakers progress through the complex set of processes that are required to produce an utterance, there are many points at which problems can occur and it can become necessary to buy some extra time.

In planning the overall message, the complexity and the length of the message can affect the likelihood of hesitation or repair. Giving instructions on how to fix the brakes on a bicycle involves mentioning more tools, more mechanical parts and more actions than simply inflating a tyre, so keeping track in memory of all the elements required to give a successful string of instructions for the former task is far more demanding, cognitively, than issuing the simpler set of instructions in the latter. Longer and more complex utterances are more prone to hesitation and repair-inducing errors.

Also at the level of overall planning, decision making can take time. For example, if a speaker needs to respond to a question by retrieving information from long term memory, then planning the response will take longer when the information is harder to retrieve (as demonstrated in responses to quiz questions: Brennan, Williams, 1995; Smith, Clark, 1993). When it is hard to plan a response, speakers will find time by using hesitation disfluencies.

Planning at the conceptual level is not only about decision making and information retrieval: It can also involve error detection and correction. During planning, a speaker may decide on a goal, begin to execute the plan and then realize that the plan is not ideal and decide to make alterations. This will entail some kind of repair, possibly correcting a factual error, or possibly modifying the output to be more accurate or specific, as suggested by Levelt's Appropriateness Repairs (Levelt, 1983). In the classification scheme outlined above, a change of plan can take the form of an insertion repair, where a slight modification is required, a substitution repair, where an alternative is preferred, or even a deletion, if the entire plan is to be ditched and a new one commenced.

Once a speaker has the overall plan ready and is accessing the lexicon, there are more challenges that can result in disfluency. Words that we learned at an earlier age and words that we use more frequently can be accessed from the lexicon more

quickly (e.g., Jescheniak, Levelt, 1994; Morrison, Ellis & Quinlan, 1992; Oldfield, Wingfield, 1965). Late-learned, less frequent words take longer to access and are therefore more likely to be preceded by hesitations.

Similarly, words with lower name agreement (more competing lexical items for the same concept) attract more hesitations (Hartsuiker, Notebaert, 2010). Physical characteristics of the words themselves can also affect their accessibility and, one can assume, their likelihood of being preceded by hesitations. Additionally, in experimental settings longer words have been found to elicit longer reaction times (e.g., Bates, D'Amico, Jacobsen, Székely, Andonova, Devescovi, Herron, Lu, Pechmann, Pléh, Wicha, Federmeier, Gerdjikova, Gutierrez, Hung, Hsu, Iyer, Kohnert, Mehotcheva, Orozco-Figueroa, Tzeng & Tzeng, 2003; Severens, Lommel, Ratinckx & Hartsuiker, 2005). So, several characteristics of lexical items can affect the likelihood that they will be preceded by hesitation.

Context can also influence hesitation before words. The context of the current discourse has an impact on the probability that a speaker will have trouble accessing a word and therefore need to hesitate. As far back as the 1950s, it was show that the probability of a word given the local context was inversely related to the probability of a hesitation preceding that word (Goldman-Eisler, 1958a; Lounsbury, 1954). Another contextual influence is that of interference with lexical access caused by the close proximity of phonologically similar words. In an experiment with running speech, words were more likely to be preceded by hesitation disfluencies when they were adjacent to phonologically similar words (e.g., "hand the hammer" Jaeger, Furth & Hilliard, 2012).

So, in addition to the intrinsic properties of lexical items, the context in which they occur can influence hesitation.

Errors in lexical access are relatively rare in typical speech, though they do occur, mostly with semantically related words, as in the example above, repeated here: "Go along the road and turn left – turn right at the traffic lights". When they do occur, they are usually repaired. With one word being substituted for another, such a repair would be classified as a substitution repair.

Once words have been accessed, phonological anticipation, perseveration and exchange errors can come about as a result of confusion in working memory. Look at the following example produced by a BBC radio reporter: "We still have this ban on cabbage – uh – on cabin baggage". Here, the words "cabin" and "baggage" had been accessed and were ready to articulate. The speaker had no plan to say "cabbage" and the semantic associations of the word were presumably not active. But confusion in working memory between the two target words caused the blend to be produced. The substitution repair followed after a brief hesitation.

Hesitation and repair disfluencies, do, of course, occur together. When an error occurs and a repair is needed, this demands some processing time. So some repairs are accompanied by hesitation phenomena (as in the cabbage/baggage example above). However, it is not the case that filled pauses regularly mark repair sites. Studies of disfluencies in speech corpora suggest that a small minority (6-10%) of

repetition and repair disfluencies are accompanied by filled pauses at the interruption point (Lickley, 1994; Nakatani, Hirshberg, 1994; Shriberg, 1994), and it is often the case that repair follows the interruption point so rapidly that there is no gap at all between interruption and repair (Blackmer, Mitton, 1991).

This brief overview of disfluency in typical speech is intended to give the reader a summary of the main forms that typical disfluencies take and an explanation of why they occur. For a more complete discussion, see Lickley (2015) and for a discussion of some phonetic aspects, see Shriberg (2001).

4. *Disfluency in stuttering*

Now, we turn to disfluencies in stuttering. The discussion will focus on the speech of adults who stutter. As proficient speakers of their native language, we should start by stating that there is no reason in principle for adults who stutter (AWS) not to produce exactly the same repertoire of typical disfluencies as anyone else. Is it simply that stuttering is at the high end of a continuum of typical disfluency with AWS producing far more frequent disfluencies than typical speakers? The answer is clearly 'No'. Stuttering is different. Not only do stuttered disfluencies have different physical and phonetic realizations from typical disfluencies, but their causes are also quite different. So, in addition to an expectation that AWS should be prone to the same types of disfluency as typical speakers, they produce additional 'stuttering-like disfluencies'.

Stuttering-like disfluencies typically occur more frequently than typical disfluencies. Wingate (2002) suggests an average of around 10 stuttering events per 100 words as a reasonable estimate of an average, taken from a range of studies. But speakers vary widely in the frequency, with rates of 50% reported in severe cases, while some cases self-report as stuttering, while apparently fluent. The frequency of stuttering also varies considerably within speaker: Some AWS report being fairly fluent on some occasions and very disfluent on others. While typical disfluency is usually rated by frequency alone, stuttering severity can also be measured by the duration of instances of stuttering. It is estimated that instances of stutters average around 1 second in duration, while rarely extending beyond 5 seconds (Bloodstein, 1944, 1987, cited in Guitar (2006)).

Disfluencies in stuttering are usually characterized as consisting of three types, according to popular academic text books and official diagnostic documents: Repetitions, Blocks and Prolongations. This is an unsatisfactory characterization, for a number of reasons, explained below.

One issue is that we need more detail than simply 'repetitions, blocks and prolongations', because the description is incomplete. We discussed repetitions and prolongations in some detail above, with reference to typical speech. But it is crucial to understand that in stuttering, these repetitions and prolongations are different. A key characteristic of stuttered disfluencies is that stuttering is usually accompanied by physical tension. This factor is ignored in that simple three-way description, yet,

it is one major factor that clearly differentiates stuttered from typical disfluency. Tension is focused on the muscle systems involved in speech articulation. Alongside the tension is the fact that stuttered disfluencies usually last longer than typical disfluencies. Whereas typical disfluencies rarely take more than a fraction of a second, stuttered disfluencies can involve the speaker getting stuck on a sound for more than one second, and sometimes several seconds. And finally, a stuttering episode may contain a mixture of blocking, prolongation and repetition.

A second issue is that the simple three-way description implies that there may be three different causes for stuttering. After all, we've described a number of different causative factors behind the various typical disfluencies. But in the case of stuttering, it seems most plausible that the three types are articulatory realizations of the same underlying difficulty: An inability to progress from one sound to the next. To some extent the realizations of this inability to progress are conditioned by the nature of the units of speech (phonemes and syllables) that are involved.

A stuttered repetition will usually involve a sound that can be repeated easily, like an oral stop consonant or cluster (electr-tr-tr-tronically), or a whole syllable (au-au-au-australia) or short word (I I I work in an office). Such repetitions are typically very rapid and may consist of just one repeated token or (more often) several repeated tokens.

A stuttered block will usually involve a long, tense articulatory closure on a stop consonant (perhaps representing a more severe case than a repetition: elect::::ronically). But blocks also occur on the glottal closure before a word-onset vowel ([ʔ]::::oil).

A stuttered prolongation usually involves a sound that can be prolonged easily, like a continuant (ssssssssseven). In considering prolongations, it is most evident that the repetition-block-prolongation distinction should be seen as realizations of the same basic phenomenon: If a speaker becomes stuck on a [t] sound, for example, and cannot release it into the following vowel, then one might class this as a 'block'; but from a functional point of view it is no different from a speaker being unable to make the transition from a [s] to a following vowel and prolonging the [s].

In all cases, if the repetition, block or prolongation leads to a successful release, then there is likely to be an intense burst, because of the extended build-up of subglottal pressure. It is also important to note that stutters mostly occur on stressed syllables, or on utterance-initial syllables (Natke, Grosser, Sandrieser & Kalveram, 2002).

This somewhat simplistic description belies the fact that close analysis and description of episodes of stuttering is usually quite problematic. For example, a repetition may vary in quality, with intermittent vocalization (t-t-t-tuh-tuh-t-t-t-take); A block may become a partially-blocked repetition (b:::::ba-b:::::ba- b::: back); a tense prolongation of /s/ may vary in intensity, giving the impression of the rapid repetition of the sound. If a successful release is not achieved, the speaker may repeat the process, with further attempts.

The picture is further complicated by the production of secondary speech behaviours. An adult with many years of experience of stuttering will typically have

learned speech behaviours that are intended to enable them to get through a stutter successfully. Amongst these are deliberately slowing down or prolonging sounds, soft articulatory contacts, resembling slurring of speech, using well-practiced fluent words or phrases intended to launch them into a fluent stream of speech (e.g., "So", "I mean", "well", "what I mean is", "Ok" …), use of monotone or unusual intonation, or intentionally repeating sounds or words in a slow and controlled manner (voluntary stuttering). All of these features make the description of stuttered speech more challenging.

In addition to the speech-based characteristics of stuttering, many AWS present with secondary behaviours, such as eye-rolling and blinking, head movements, tongue thrusts, facial grimaces, movements of arms or hands.

In general, there is a lot of interspeaker variation in the presentation of stuttering, and perhaps because of that, in addition to the complexity outlined above, a lack of detailed phonetic description of stuttered speech.

What can be said is that for the most part, stuttering-like disfluencies are different in form from typical disfluencies. While words and syllable onsets are repeated in typical speech, it is relatively rare for a word or sound to be repeated more than once. While there is prolongation in typical speech, the duration of prolongations is relatively short, where it can extend over a second in duration in stuttering, and it is more common at the ends of words in typical speech, rather than at the syllable onsets that are affected by stuttering. Blocking is not considered by researchers examining typical disfluency. There is no evidence of physical tension associated with typical disfluency, but this is a defining characteristic of stuttering-like disfluency.

Earlier, we discussed how typical disfluencies can be explained with reference to the various levels of processing involved in speech production. Repetitions and prolongations are found frequently in typical speech, and we have attributed them to problems at the conceptualization and formulation stages, where the speaker needs to buy time. But in stuttering, speakers report that when they get stuck on a word, they know exactly what they want to say, so there is no problem with conceptualization or formulation. Although there is some evidence that word-finding difficulties may also play a role (Hubbard, Prins, 1994; Prins, Main & Wampler, 1997), the main problem lies in executing the plan for articulation. While it is tempting to say that a speaker who stutters on a given sound has a problem with that sound, this is a mischaracterization. The problem lies in the transition from the repeated, blocked or prolonged sound to the next sound, in synchronizing the change in articulation and/or voicing between one sound and the next. This view forms the basis of Wingate's fault-line hypothesis (Wingate, 1988), which suggests that in stuttering a speaker suffers a delay in the syllable rhyme following production of the onset. A similar view is also modelled in the EXPLAN Theory (Howell, AuYeung, 2002), which posits that stuttering may occur when there is a misalignment between a plan and execution such that a syllable is initiated before the plan for the rest of the syllable is ready.

Beyond the stuttered disfluencies themselves, there are other factors related to stuttering that can potentially result in disfluencies that may be seen as more typical. Adults who stutter are very aware that they are likely to stutter in certain situations and on certain words and sounds, and this creates apprehension (Blood, Blood, Tellis & Gabel, 2001; Kelso, 1997; Neiman, Rubin, 1991).

In his Anticipatory Struggle Hypothesis, Bloodstein (1975) viewed the anticipation of difficulty in speaking as a cause of the muscular tension and then the fragmentation of speech that is characteristic of stuttering. Although there is little support for theories of stuttering that place anxiety about speech at the centre, there is some evidence that anticipation of trouble in speech can increase the likelihood that an AWS will stutter on a word (Brocklehurst, Lickley & Corley, 2012). It has been demonstrated that AWS are more sensitive than are typical speakers to minor disfluencies in the speech of other people, as well as in their own speech (Lickley, Hartsuiker, Corley, Russell & Nelson, 2005). But often, AWS plan ahead to avoid the possibility of stuttering. In some cases, this may entail avoiding speaking altogether. In other cases, an AWS may make an alteration to their plan, either at the conceptual level, by deciding to change the message that they were originally going to produce, or at the lexical level, by finding an alternative word, to avoid one that they anticipate to present difficulties. These processes can take time, and may result in hesitation or reformulation of an utterance, creating disfluency, though potentially avoiding stuttering. On the other hand, some AWS are so adept at circumlocution that they appear fluent, although they are still well aware of the underlying problem: This is usually referred to as covert stuttering (Murphy, Quesal & Gulker, 2007). It is not clear whether a person with covert stuttering can be so controlled in their speech that they avoid typical disfluencies, too, though it is likely that they would present with an elevated rate of hesitations.

5. *Conclusion*

It is clear, then, that stuttered disfluencies differ greatly in form from typical disfluencies. Although similar terms have been used to describe them (repetitions, prolongations, silent pauses or blocks), they differ in quantity as well as in quality. In more severe cases, stutters occur at a far greater frequency than typical disfluencies. The more prominent typical disfluencies are relatively short (just over 400 ms for the filled pause "um", Lickley, 2015), whereas stutters can average around 1 second in duration and extend to 5 seconds or more. Repetitions in stuttering involve multiple repeats of a sound or syllable, whereas it is relatively rare for a typical disfluent repetition to consist of more than one repeat. Similarly, stuttered prolongations are typically significantly longer than those in typical speech. Stuttering-like blocks are extremely rare in typical speech, though brief silent pauses are not uncommon. Finally, stuttering-like disfluencies are typically accompanied by muscular tension, which is hardly ever present in typical disfluency.

The fact that stuttering-like disfluencies are different from typical disfluencies in quality and quantity does not necessarily rule out the possibility that they are just extreme cases of certain types of otherwise typical hesitation disfluencies. However, it is also clear that stuttered disfluencies differ greatly from typical disfluencies in their provenance. In most cases, typical disfluencies are responses to difficulties in planning and formulation of speech. In stuttering, the primary symptoms result from difficulties in the coordination of motor programmes for the execution of existing plans. Further evidence for this come from the observation that in cases where the planning of the message and the words to be used are already supplied – in reading out loud – typical speakers are only rarely disfluent, whereas many AWS stutter at a higher rate than in spontaneous speech.

We began with an operational definition of fluency. We then discussed how between typical speakers and speakers who stutter the speech of the two populations is characterized by different types of disfluency. Some researchers object to the term disfluency being used to describe features of speech that is entirely typical. In this sense, for some adults who stutter, a therapy goal might be 'typical fluency' insofar as that alternative definition of 'fluency' might include typical disfluency.

Bibliography

ANDERSON, A.H., BADER, M., BARD, E.G., BOYLE, E., DOHERTY, G., GARROD, S., ISARD, S., KOWTKO, J., MCALLISTER, J., MILLER, J., SOTILLO, C., THOMPSON, H.S. & WEINERT, R. (1991). The HCRC Map Task Corpus. In *Language and Speech*, 34(4), 351-366.

BATES, E., D'AMICO, S., JACOBSEN, T., SZÉKELY, A., ANDONOVA, E., DEVESCOVI, A., HERRON, D., LU, C.C., PECHMANN, T., PLÉH, C., WICHA, N., FEDERMEIER, K., GERDJIKOVA, I., GUTIERREZ, G., HUNG, D., HSU, J., IYER, G., KOHNERT, K., MEHOTCHEVA, T., OROZCO-FIGUEROA, A., TZENG, A. & TZENG, O. (2003). Timed picture naming in seven languages. In *Psychonomic Bulletin & Review*, 10(2), 344-380.

BELL, A., JURAFSKY, D., FOSLER-LUSSIER, E., GIRAND, C., GREGORY, M. & GILDEA, D. (2003). Effects of disfluencies, predictability, and utterance position on word form variation in English conversation. In *The Journal of the Acoustical Society of America*, 113, 1001-1024.

BLACKMER, E.R., MITTON, J.L. (1991). Theories of monitoring and the timing of repairs in spontaneous speech. In *Cognition*, 39, 173-194.

BLOOD, G.W., BLOOD, I.M., TELLIS, G. & GABEL, R. (2001). Communication apprehension and self-perceived communication competence in adolescents who stutter. In *Journal of Fluency Disorders*, 26(3), 161-178.

BLOODSTEIN, O. (1975). Stuttering as tension and fragmentation. In EISENSON, J. (Ed.), *Stuttering: A second symposium*. New York: Harper & Row, 1-96.

BORTFELD, H., LEON, S.D., BLOOM, J.E., SCHOBER, M.F. & BRENNAN, S.E. (2001). Disfluency rates in conversation: Effects of age, relationship, topic, role, and gender. In *Language and Speech*, 44(2), 123-147.

Brennan, S.E., Williams, M. (1995). The feeling of another's knowing: Prosody and filled pauses as cues to listeners about the metacognitive states of speakers. In *Journal of Memory and Language*, 34(3), 383-398.

Brocklehurst, P.H., Lickley, R.J. & Corley, M. (2012). The influence of anticipation of word misrecognition on the likelihood of stuttering. In *Journal of Communication Disorders*, 45(3), 147-160.

Butcher, A. (1981). Aspects of the speech pause: Phonetic correlates and communicative functions. PhD Dissertation, Christian Albrechts Universität zu Kiel.

Clark, H.H., Wasow, T. (1998). Repeating words in spontaneous speech. In *Cognitive Psychology*, 37(3), 201-242.

Duez, D. (1993). Acoustic correlates of subjective pauses. In *Journal of Psycholinguistic Research*, 22(1), 21-39.

Eklund, R. (2004). Disfluency in Swedish human-human and human-machine travel booking dialogues. PhD Dissertation, University of Linköping.

Ferreira, F. (1993). Creation of prosody during sentence production. In *Psychological Review*, 100, 233-253.

Ferreira, F. (2007). Prosody and performance in language production. In *Language and Cognitive Processes*, 22(8), 1151-1177.

Fox Tree, J.E. (1995). The effects of false starts and repetitions on the processing of subsequent words in spontaneous speech. In *Journal of Memory and Language*, 34, 709-738.

Fox, B.A., Jasperson, R. (1995). A syntactic exploration of repair in English conversation. In Davis, P.W. (Ed.), *Descriptive and theoretical modes in the alternative linguistics*. Amsterdam: John Benjamins, 77-134.

Goldman-Eisler, F. (1958a). The predictability of words in context and the length of pauses in speech. In *Language and Speech*, 1(3), 226-231.

Goldman-Eisler, F. (1958b). Speech production and the predictability of words in context. In *Quarterly Journal of Experimental Psychology*, 10(2), 96-106.

Goldman-Eisler, F. (1961). A comparative study of two hesitation phenomena. In *Language and Speech*, 4(1), 18-26.

Grosjean, F., Collins, M. (1979). Breathing, pausing and reading. In *Phonetica*, 36(2), 98-114.

Guitar, B. (2006). *Stuttering: An integrated approach to its nature and treatment* (3rd ed.). Philadelphia: Lippincott Williams and Wilkins.

Hartsuiker, R.J., Notebaert, L. (2010). Lexical access problems lead to disfluencies in speech. In *Experimental Psychology (Formerly Zeitschrift Für Experimentelle Psychologie)*, 57(3), 169-177.

Heeman, P.A. (1997). Speech repairs, intonational boundaries and discours markers: Modeling speakers' utterances in spoken dialog. PhD Dissertation, University of Rochester.

Howell, P., Au-Yeung, J. (2002). The EXPLAN theory of fluency control applied to the diagnosis of stuttering. In *Amsterdam Studies in the Theory and History of Linguistic Science Series*, 4, 75-94.

Hubbard, C.P., Prins, D. (1994). Word familiarity, syllabic stress pattern, and stuttering. In *Journal of Speech, Language, and Hearing Research*, 37(3), 564-571.

Jaeger, T.F., Furth, K. & Hilliard, C. (2012). Phonological overlap affects lexical selection during sentence production. In *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 38(5), 1439-1449.

Jescheniak, J.D., Levelt, W.J. (1994). Word frequency effects in speech production: Retrieval of syntactic information and of phonological form. In *Journal of Experimental Psychology-Learning Memory and Cognition*, 20(4), 824-843.

Kelso, K. (1997). The relationship between communication apprehension, communication competence, and locus of control in stutterers and nonstutterers. In *Journal of Fluency Disorders*, 2(22), 153.

Levelt, W.J.M. (1983). Monitoring and self-repair in speech. In *Cognition*, 14, 41-104.

Lickley, R.J. (1994). Detecting disfluency in spontaneous speech. PhD Dissertation, University of Edinburgh.

Lickley, R.J. (1998). *HCRC disfluency coding manual.* HCRC Technical Report No. TR-100. University of Edinburgh.

Lickley, R.J. (2001). Dialogue moves and disfluency rates. In *Proceedings of DiSS: ISCA Tutorial and Research Workshop on Disfluency in spontaneous speech*, Edinburgh, UK.

Lickley, R.J. (2015). Fluency and disfluency. In Redford, M. (Ed.), *The handbook of speech production.* Wiley-Blackwell, 445-469.

Lickley, R.J., Hartsuiker, R.J., Corley, M., Russell, M. & Nelson, R. (2005). Judgment of disfluency in people who stutter and people who do not stutter: Results from magnitude estimation. In *Language and Speech*, 48(3), 299-312.

Lounsbury, F.G. (1954). Transitional probability, linguistic structure and systems of habit-family hierarchies. In Osgood, C.E., Sebeok, T.A. (Eds.), *Psycholinguistics: A survey of theory and research problems*. Baltimore: Williams and Wilkins, 93-101.

Maekawa, K. (2004). Design, compliation and some preliminary analyses of the corpus of spontaneous Japanese. In Maekawa, K.Y.K. (Ed.), *Spontaneous speech: Data and analysis*. Tokyo: The National Institute for Japanese Language, 87-108.

Morrison, C.M., Ellis, A.W. & Quinlan, P.T. (1992). Age of acquisition, not word frequency, affects object naming, not object recognition. In *Memory & Cognition*, 20(6), 705-714.

Murphy, B., Quesal, R.W. & Gulker, H. (2007). Covert stuttering. In *Perspectives in Fluency and Fluency Disorders*, 17(2), 4-9.

Nakatani, C.H., Hirschberg, J. (1994). A corpus-based study of repair cues in spontaneous speech. In *The Journal of the Acoustical Society of America*, 95, 1603-1616.

Natke, U., Grosser, J., Sandrieser, P. & Kalveram, K.T. (2002). The duration component of the stress effect in stuttering. In *Journal of Fluency Disorders*, 27(4), 305-318.

Neiman, G.S., Rubin, R.B. (1991). Changes in communication apprehension, satisfaction, and competence in foreign dialect and stuttering clients. In *Journal of Communication Disorders*, 24, 353-366.

Oldfield, R.C., Wingfield, A. (1965). Response latencies in naming objects. In *Quarterly Journal of Experimental Psychology*, 17(4), 273-281.

Oviatt, S. (1995). Predicting spoken disfluencies during human-computer interaction. In *Computer Speech and Language*, 9, 19-35.

PLAUCHÉ, M., SHRIBERG, E. (1999). Data-driven subclassification of disfluent repetitions based on prosodic features. In *Proceedings of the International Congress of Phonetic Sciences*. San Francisco, USA, 1513-1516.

POSTMA, A., KOLK, H. (1993). The Covert Repair Hypothesis: Prearticulatory repair processes in normal and stuttered disfluencies. In *Journal of Speech and Hearing Research*, 36, 472-487.

PRINS, D., MAIN, V. & WAMPLER, S. (1997). Lexicalization in adults who stutter. In *Journal of Speech, Language, and Hearing Research*, 40 (2), 373-384.

SEVERENS, E., LOMMEL, S.V., RATINCKX, E. & HARTSUIKER, R.J. (2005). Timed picture naming norms for 590 pictures in Dutch. In *Acta Psychologica*, 119 (2), 159-187.

SHRIBERG, E. (1994). Preliminaries to a theory of speech disfluencies. PhD Dissertation, University of California.

SHRIBERG, E. (2001). To 'errrr' is human: Ecology and acoustics of speech disfluencies. In *Journal of the International Phonetic Association*, 31(1), 153-169.

SMITH, V.L., CLARK, H.H. (1993). On the course of answering questions. In *Journal of Memory and Language*, 32(1), 25-38.

WANG, Y.T., GREEN, J.R., NIP, I.S., KENT, R.D. & KENT, J.F. (2010). Breath group analysis for reading and spontaneous speech in healthy adults. In *Folia Phoniatrica Et Logopaedica*, 62(6), 297-302.

WATANABE, M. (2009). *Features and roles of filled pauses in speech communication: A corpus-based study of spontaneous speech*. Tokyo: Hituzi Syobo Publishing.

WINGATE, M.E. (1988). *The structure of stuttering*. New York: Springer.

WINGATE, M.E. (2002). *Foundations of stuttering*. San Diego, CA: Academic Press.