The Influence of Lexical, Conceptual and Planning Based Factors on Disfluency Production

Michael J. Schnadt (m.j.schnadt@sms.ed.ac.uk)
School of Philosophy, Psychology and Language Sciences,
Edinburgh University, 7 George Square, Edinburgh, EH8 9JZ, UK

Martin Corley (martin.corley@ed.ac.uk)
School of Philosophy, Psychology and Language Sciences,
Edinburgh University, 7 George Square, Edinburgh, EH8 9JZ, UK

Abstract

Two experiments were conducted to elicit naturalistic speech, while manipulating factors thought to influence disfluency production. Participants described the route taken by a marker through visually presented networks of objects linked via one or more paths. In Experiment 1, lexical frequency and name agreement of the object names were manipulated; in Experiment 2, linguistic properties were kept constant and accessibility was manipulated through visual blurring. An increase in disfluency was observed immediately preceding object names in cases where the objects named were either low frequency or blurred. In both experiments, prolongations were the most frequently occurring class of disfluency. Additionally, when disfluencies during the path description were examined, more possible path choices led to greater numbers of disfluencies, which were predominantly filled pauses and repairs. This study allows us to draw preliminary conclusions about the influence of lexical, conceptual and planning-based factors on disfluency production and to begin to determine precisely the circumstances under which disfluencies occur in natural speech.

Introduction

Natural spoken language is full of disfluency, generally defined as "phenomena that interrupt the flow of speech and do not add propositional content to an utterance" (Fox Tree, 1995). These phenomena include pauses, interruptions, substitutions, repeated words or phrases, prolongations, such as *the* pronounced "thee", and filled pauses such as *um* and *uh*. Such disruptions are very frequent: Brennan and Schober (2001) estimate that around 10% of utterances contain at least one disfluency, whereas Fox Tree (1995), averaging across a number of studies, estimated that the rate of disfluencies in spontaneous speech is about 6 per 100 words.

While the distribution of disfluencies in spontaneous speech is relatively well understood from corpus-based studies (e.g., Bortfield, Leon, Bloom, Schober & Brennan, 2001; Clark & Fox Tree, 2002; Shriberg, 1996), and claims have been made as to the differing functions of different types of disfluency (e.g., Clark & Fox Tree, 2002), experimental studies to date have proved somewhat inconclusive as to their cause. Take, for example, the finding that disfluencies are more likely to precede open-class words (Maclay & Osgood, 1959). On current evidence, it is entirely unclear what the underlying cause of these disfluencies might be. They could be a consequence of the relatively low frequencies (compared to closed-class words) with which open-class words are likely to occur (Levelt, 1983), or in other words, they could be caused by lexical retrieval difficulties. Alternatively, disfluencies could be attributed to the vastly greater choice of open-class words available to the speaker (Schachter, Christenfeld, Ravina & Bilous, 1991), or to difficulties with lexical choice and access. Finally, they may be due to causes outwith the language system: if, for example, a speaker is trying to name an unfamiliar or ambiguous object (Siegman & Pope, 1966). Effectively, the difficulties signaled by disfluencies could occur at any stage of the speech process: during planning, lexical retrieval or articulation of the speech plan, and it has been argued that different types of disfluency may signal different kinds of problems during production (Bortfield et al., 2001). Clearly, a better understanding of the underlying causes of disfluency would provide an important contribution our understanding of language production in general.

In this paper, we present two experiments which use the Network Task (Levelt, 1983; Oomen & Postma, 2001) to explicitly manipulate the content of what people say when describing a network of objects, in order to investigate a priori what factors influence the production of disfluency. In Experiment 1, the frequency and name agreement of items in the networks are varied; Experiment 2 varies the visual accessibility of pictures used, reflecting difficulties that do not have their origin in the linguistic system. These experiments allow us to investigate the causes of disfluency directly, establishing whether different disfluencies serve different purposes.

Experiment 1

Method

The experiment was presented as a communication task. Participants described the route taken by a marker through a network of objects to a listener situated behind a screen.

Participants Twenty students from the University of Edinburgh participated in the experiment (8 male, 12 female). All were native British English speakers.

Materials The experimental materials used in the Network Task consisted of 6 visually presented networks based on those of Oomen and Postma (2001). Each network contained 8 black and white line drawings, arranged in different configurations and connected by one, two or three straight or curved lines. Each network was associated with a route through the objects. A route always consisted of nine steps, where six of the pictures occurred once in the route, and two of the pictures occurred twice. The route through the network that participants had to follow in their descriptions was indicated by a red marker that moved along the lines connecting each picture (see Figure 1 for an example).

To construct the networks, 36 experimental and 12 filler pictures were chosen from Snodgrass and Vanderwart's (1980) picture set. Objects were selected based on name agreement scores taken from Griffin and Huitema (1999). Included objects had either at least 90% agreement or less than 50% agreement with the target name. Only pictures that were given alternative names due to intrinsic differences in the speakers' own language processes and representations were included. Objects given alternative names due to picture ambiguity were excluded.

Lemmatized word frequencies were taken from Kucera and Francis (1967). Where a word did not appear in Kucera and Francis (1967), the British National Corpus (BNC) frequency list (Kilgariff, 1995) was used, and the frequency converted to an equivalent scale. 18 high frequency words with frequencies ranging from 19 (guitar) to 591 (computer) per 1.5 M (mean = 132.9) with high name agreement and 18 low frequency words with frequencies ranging from 0.26 (gavel) to 20 (suitcase) per 1.5M (mean = 5.4) with low name agreement were selected.

Individual networks comprised 3 pictures with names of high frequency and high name agreement (hereafter referred to as HF) and 3 pictures with names of low frequency and low name agreement (referred to as LF). Two fillers were also included at nodes that the marker passed through twice. Each network was constructed so that the picture sequences were alternated between HF and LF and so that semantically and phonologically related targets were not adjacent.

Participants' speech was recorded on a SONY TLD-D8 DAT Walkman recorder using a Sennheiser C6 microphone. The networks were presented sequentially on a stand-alone computer screen, 40cm away from the participant.

Procedure The speaker and confederate listener sat on opposite sides of a screen. While both speaker and listener could hear each other clearly, the screen occluded any visual interaction.

Participants were informed that the experiment investigated how well individuals followed instructions without the aid of eye contact or body language. They would see a series of networks and it was their task to describe the route of a marker through each network so that the listener could fill in the route and picture names on a blank network. Participants were instructed to describe the path of the marker, making sure to include the name of each picture, the direction the marker was taking, the angle and shape of the line the marker was moving along, and the position of the line in relation to the others. The speaker was instructed to talk in complete sentences and to modulate their speech rate to keep up with the position of the marker (as in Oomen & Postma, 2001). Participants were told that they had to be as accurate as possible as the listener could not ask for any repeated descriptions.

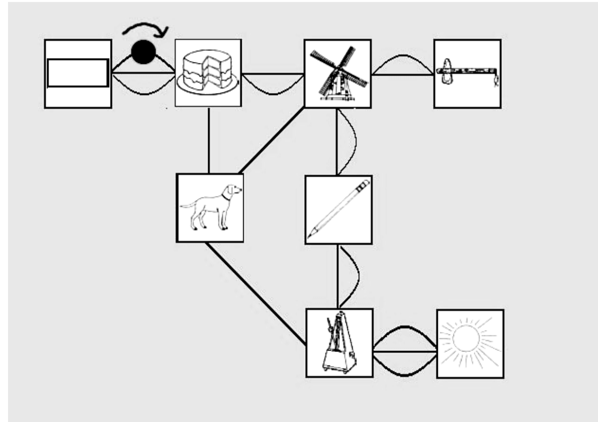

Figure 1: An example network.

The order of network presentation was randomized across participants. The marker started moving through the first network when the spacebar on the keyboard was pressed. It took 30 seconds to pass through the whole network, which was found to be an appropriate speed for eliciting fast connected spontaneous speech containing sufficient numbers of disfluencies. Upon completion, participants pressed the spacebar again to reveal the next network.

Transcription and Coding Transcription and coding was carried out by the first author. 20% of the transcriptions were independently verified and coded by another researcher from the Edinburgh Disfluency Group. The two coders agreed on 84% of identified disfluencies. Items for which transcription or coding differed were re-examined by both researchers until a consensus was reached.

The analysis presented here is based on three classes of disfluency: filled pauses (such as *um* or *uh*); prolongations (such as *ay*, *thee*, or prolonged *to*); and repairs, including repetitions and corrections (e.g., insertions, substitutions or deletions). These are illustrated in Table 1, with examples.

Results

The mean (and standard deviation) number of alternative names used for each picture was 1.50 (.71) for HF items and 5.28 (3.12) for LF items. As expected, LF targets had significantly more names used for them than HF targets (t (19) = 5.01, $p < .001$).

Table 1: Coding scheme for disfluencies.

Disfluency Class	Disfluency type	Example (in bold)
Filled Pause	Um	"Right to the_**um** trophy along the straight line." (subject 4, network 5)
	Uh	"…it's going up to the **uh** house…" (subject 2, network 6)
Prolongation	Prolongation - *Ay*	"Then the dot's moving left to **ay** table on a left straight line" (subject 5, network 5)
	Prolongation – *Th:e*	"…down to **th:e** lion badge…" (subject 8, network 6)
	Prolongation – *to:o*	"…along **to:o** a hammer."(subject 2, network 5)
Repair	Repeat	"…towards the **s_s**ickle um along the…" (subject 6, network 4)
	Correction	"…towards the **s_rectangle**" (subject 20, network 2)

Out of a total of 720 items, 29 items were not named in participants' descriptions (7 HF and 22 LF items), and were removed from further analysis. Disfluency analysis was performed from the remaining 691 data points.

Table 2 shows the percentage of utterances that contained a disfluency overall, percentages by disfluency class and for information, by individual disfluency types. It also details the proportional increase in disfluencies produced between the HF and LF conditions. A 1-way ANOVA on disfluencies immediately preceding the target item showed a significant overall effect of item type (LF vs. HF) on the total number of disfluencies produced (F_1 (1,19) = 22.01, $p < .001$; F_2 (1,34) = 24.13, $p < .001$). When disfluency classes were examined, LF utterances contained significantly more prolongations in the Target position than did HF utterances (F_1 (1, 19) = 25.33, $p < .001$; F_2 (1,34) = 24.83, $p < .001$). Filled pauses were present only marginally more often in the LF condition (F_1 (1, 19) = 3.23, $p = .088$), although this was significant by items (F_2 (1,34) = 6.43, p = .016). It is likely that this marginal result is due to the low occurrence of filled pauses preceding target items in the data set (27 fillers were observed in this position in total, 7 HF and 20 LF, compared to 160 prolongations). The class of repairs showed no significant difference between LF and HF items (F_1, F_2 both < 1)[1].

Table 2: Proportion of utterances containing disfluency

Disfluency Class	HF	LF	Proportional Increase
Total Disfluency	19.2%	36.9%	1.92
Filled Pause	2.0%	5.6%	2.8
Um	1.7%	3.4%	2
Uh	0.3%	2.6%	8.67
Prolongation	12.5%	32.5%	2.6
Repair	2.7%	3.9%	1.44
Repetition	*1.2%*	*2.1%*	*1.71*
Correction	*1.5%*	*1.9%*	*1.27*

[1] Individual disfluency types were also examined, but no significant effects were found for types of filled pause or repairs, with the exception of *uh*, which showed an effect that approached significance (F_1 (1, 19) = 3.85, $p = .065$; F_2 (1,34) = 5.88, $p = .021$), but was based on only 10 data points.

Discussion

The purpose of Experiment 1 was to examine the influence of lexical frequency of an object's name on the production rate of disfluencies immediately preceding naming of a target item. The results demonstrate that pictures possessing names of low frequency and low name agreement yielded significantly more disfluencies than for object names of high frequency and high name agreement. Thus, difficulties associated with lexical access and retrieval of word forms do appear to be accommodated through the use of disfluencies.

When disfluencies were broken down into classes, only the number of prolongations produced showed a significant increase preceding low frequency items. Additionally, prolongations occurred much more frequently preceding the target item than did other classes of disfluency. Prolongations comprised over 75% of observed disfluencies, compared to about 13% and 10%, respectively, for filled pauses and repairs. Perhaps surprisingly, given that much of the literature concerns them, very few filled pauses were observed preceding target names, although the proportionate increase was similar to that for prolongations. This may have been due to the fact that participants were instructed to modulate their speech rate to the pace of the marker, and under this time-pressured condition were less likely to use a filler term.

The relative abundance of prolongations in the Target position suggests that they are specifically used to accommodate short term difficulties in the retrieval of word forms, and that these difficulties are resolved more quickly than those that require the formulation and insertion of a filler term into upcoming speech. In comparison, Clark & Fox Tree (2002) have proposed that filled pauses are used to signal a variety of production difficulties, including lexical access and retrieval.

Experiment 2

Experiment 2 sought to examine problems associated with the conceptual formulation of the speech message. This was performed by blurring the object images, which made identification of the items and therefore conceptual generation of a message more difficult.

Method

The experimental method was identical to that of Experiment 1, with the exception of the materials used in the networks, as described below.

Participants 20 students (7 male, 13 female) from the University of Edinburgh participated in the experiment. All were native British English speakers.

Materials 8 visually presented networks were used, based on those of Oomen and Postma (2001), each containing 8 pictures of objects. 64 objects were selected from the Snodgrass and Vanderwart (1980) set of 260 objects. The images used were grayscale textured images, which were found to be more suitable for blurring (taken from Rossion & Portois, 2004).

All selected images had Kucera and Francis (1967) frequency counts of 10-30 per 1.5 million (mean = 19), and a CELEX (Baayen, Piepenbrock, & van Rijn, 1993) frequency within the range of 8-45 counts per million (mean = 19.4). Image names were rated 4 or above on a scale of 1-5 in a rated picture name agreement measure (Barry, Morrison, & Ellis, 1997).

All of the test images were blurred using a Gaussian blur with a 1.5 pixel radius to create a set of 48 clear images and a set of 48 blurred images that had a similar degree of visual naming difficulty. 16 filler items were also selected from the same item set.

Each network contained 8 pictures (6 test items and 2 fillers), as described previously. 2 sets of networks were created with alternating clear and blurred items, so that the use of clear and blurred images was counterbalanced across participants. Items were positioned so that semantically and phonologically related items were not adjacent to each other.

Transcription and Coding Transcription and coding was carried out by the first author, and 20% of transcriptions were independently verified and coded as in Experiment 1 (mean inter-rater coding agreement was 87%).

Results

23 items (7 clear and 16 blurred) were not named and the analysis was performed from the remaining 937 utterances.

Table 3 shows the percentage of utterances that contained a disfluency, and the proportional increase in disfluencies between clear and blurred conditions. A 1-way ANOVA (blurred vs. clear) on disfluencies immediately preceding target items demonstrated a main effect of item type (F_1 (1,19) = 5.6, p = .029; F_2 (1,47) = 4.72, p = .035). When disfluency classes were examined, blurred items were preceded by significantly more prolongations than clear items (F_1 (1,19) = 7.14, p = .015; F_2 (1,47) = 6.68, p = .013). For both filled pauses and repairs these differences were non-significant (all p's > .1). Analyses of individual disfluency types for filled pauses and repairs showed no significant differences (all p's > .1).

Table 3: Proportion of items preceded by disfluency

Disfluency Class	Clear	Blurred	Proportional Increase
Total Disfluency	18.1%	24.7%	1.36
Filled Pause	3.3%	2.1%	0.64
Um	1.7%	1.7%	1.00
Uh	1.6%	0.5%	0.31
Prolongation	14.2%	21.8%	1.54
Repair	1.6%	1.9%	1.19
Repetition	0.6%	0.8%	1.33
Correction	0.8%	1.3%	162

Discussion

The results of Experiment 2 demonstrated that impeding conceptual access of object representations through image blurring did significantly affect the overall rate of disfluency production. The data for individual disfluency classes showed a similar pattern to those of Experiment 1, with prolongations comprising the majority of disfluencies observed preceding the target name, while filled pauses and repairs were observed in less than 4% of utterances.

Interestingly, the relative increase in disfluency rates between item conditions was not as great in Experiment 2 as that observed in Experiment 1, thus speakers were actually more fluent when naming blurred items in their descriptions than they were when naming low frequency items. This could be taken as evidence that as conceptualization occurs earlier in the production process than generation of the word form, speakers identify conceptual difficulties earlier, providing them with more opportunity for fluent resolution. However, since the frequency and blurring manipulations influence different processes, further research is needed to determine whether the effect of one manipulation was of a greater magnitude than the other.

Disfluencies Produced During Path Description

While access and retrieval of word forms can produce processing difficulties that result in disfluent speech, arguably the largest processing task associated with utterance generation is the formulation of the speech plan. Disfluencies and speech errors have long been studied as useful tools to gain more insight into the cognitive processes of speech planning as they are closely associated with increased processing loads on parts of the production system (Shriberg, 1996).

One aspect of the current experiment that is directly related to utterance planning is speakers' description of the path that the marker takes from one node to the next. The structure of the networks lends itself well to further analysis of these path descriptions as each pair of nodes either has 1, 2 or 3 paths connecting them, and the marker always takes one of these to pass from one node to the next. As the speaker had to describe this path at speed to keep up with the marker, multiple possible paths between nodes should increase the processing load substantially, and lead to more disfluencies. It is of interest here to examine exactly how the pattern of disfluencies produced during the path description compares to the pattern observed immediately preceding the target name. In addition, the pattern of disfluencies produced at the start of each utterance can be analyzed to see whether they are influenced by planning load or retrieval difficulty, which would indicate that they are associated with utterance planning, or whether they are being used as structural devices by the speaker.

Path Analysis & Results

Analysis of the data from Experiment 2 (which was the only suitably counterbalanced experiment) was performed on test items that were connected by 1, 2 or 3 paths from the previous item. As the path description tended to be considerably longer than the target description, it was possible that more than one disfluency, sometimes of the same type, would occur in a single utterance. Therefore, the numbers given are the average number of disfluencies per utterance as opposed to the percentage of utterances containing a particular disfluency, as reported previously.

A 2-way ANOVA (item type vs. number of paths) on the data from Experiment 2 showed a significant effect of the number of paths on the average number of disfluencies produced in each path description, both by subjects and by items, (F_1 (2, 38) = 10.6, $p < .01$; F_2 (2, 37) = 7.74, $p <$ .005). There was no effect of item type on the average number of disfluencies in the path descriptions (F_1, $F_2 < .1$, p's > .5). When each disfluency class was examined there was an effect of number of paths on filled pauses produced (by items, F_2 (2, 37) = 5.065, $p < .01$, marginal by subjects, F_1 (2, 38) = 3.159, $p = .054$). There was also a significant effect on the number of prolongations produced (F_1 (2, 38) = 10.717, $p < .001$; F_2 (2, 37) = 5.122, $p = .011$). Finally there was a main effect of number of paths on the repairs produced (F_1 (2, 38) = 5.19, $p = .01$; F_2 (2, 37) = 3.35, $p =$ .046). When individual disfluency types were examined, *um*s showed a marginally significant increase with number of paths (F_1 (2, 38) = 2.89, $p = .068$; F_2 (2, 37) = 4.83, $p =$ .014), and a significant effect was also observed for the number of corrections (F_1 (2, 38) = 4.87, $p = .013$; F_2 (2, 37) = 3.64, $p = .036$). No significant effects were observed ofr *uh*s or repetitions (p's > .2). These results are summarized in Table 4.

Across all disfluency classes, item type (clear vs. blurred) again did not have any significant effect on the number of disfluencies produced in the description of the path leading up to that item (all F's < 3.366, p's > 0.05).

Disfluencies produced during the part of the utterance classified as "beginning", i.e. before the path description had been initiated, were also analyzed. On average, 7.9% of utterances contained at least one disfluency in this position. 85% of disfluencies produced were filled pauses, of which over 70% were *um*s. A 2-way ANOVA (item type vs. number of paths) showed a marginal effect of the number of paths (F (2, 38) = 2.796, $p = .074$) and no effect of item type (F (1, 19) = .368, $p > .5$) on the number of disfluencies produced at the beginning of the utterance.

Table 4: Number of Disfluencies Per Utterance in Path Description

	1 path	**2 paths**	**3 paths**
Total	0.048	0.172	0.243
Filled Pauses	0.021	0.047	0.097
Um	0.007	0.030	0.059
Uh	0.014	0.023	0.039
Prolongations	0.000	0.032	0.054
Repairs	0.027	0.092	0.093
Repetition	0.007	0.018	0.012
Correction	0.021	0.083	0.086

Discussion

This analysis demonstrates an increase in overall disfluency production when speakers had more possible paths to choose from in their description of the marker's route. Additionally, disfluencies produced were not related to the upcoming target item, but to difficulties linked with the path description itself, suggesting that having multiple possible paths increases the processing load associated with planning and formulating the correct descriptive utterance for a given path.

Interestingly, the pattern of disfluencies observed in the path description was markedly different to that found preceding the target item. In the path description, filled pauses and repairs were more frequent than prolongations, whereas preceding the target item, prolongations occurred much more frequently than either filled pauses or repairs. The predominance of filled pauses could reflect more serious difficulties associated with utterance formulation, as opposed to access and retrieval, particularly as the words used in the path descriptions often repeated between items. Additionally, the increased processing load resulting from having to accommodate more than one path in the description is likely to have an impact on utterance planning, and this is borne out by a higher rate of observed repairs, which were primarily corrections. Due to the speeded nature of the task, speakers may have had to initiate planning of the next utterance before they were certain of where the marker was moving, and only identified their error after commencing articulation, thus requiring a repair.

The analysis of disfluencies occurring at the beginning of the utterance showed no effect of either number of paths or the item condition, indicating that disfluencies occurring here do so largely irrelevant of potential upcoming difficulties. However, the vast majority of these were filled pauses, and in particular *um*'s, which it is suggested are used as inter-phrasal continuation devices (Clark & Fox Tree, 2002; Swerts, 1998), and as such, may have been inserted either to maintain the pace of speech with the marker or as a dialogue structuring device.

General Discussion

The two experiments presented here demonstrate that factors influencing disfluency production can be manipulated using an online task where participants produce spontaneous, task-specific utterances. We examined how changes in the representational accessibility of a target item influenced disfluency production immediately preceding the item in the speaker's descriptive utterances.

The similarity of the pattern of results for disfluency classes between the two experiments gives an indication that the source of the problem within the production system, whether conceptually or lexically driven, does not appear to affect the type of disfluency used to accommodate it. It is probable that the nature of the difficulty (whether related to planning, formulation or access and retrieval), the point at which it has been identified in the production of an utterance and where the disfluency is placed within the syntactic frame of the utterance are stronger influences on the type of disfluency produced. Consequently, the prevalence of prolongations preceding the target item is likely to be due to the fact that difficulties here tend to be related to access and retrieval of the word form, which can be quickly resolved, and such items are normally preceded by easily prolonged articles like *the* or *to*. Because the disfluency is produced in or immediately preceding the noun phrase, prolongation of *the* or *to* can provide the time required to resolve the difficulty without additional articulation and in a way that maintains the integrity of the speech flow. This raises further questions about how the type of disfluency is related to where it is produced within the utterance, and how factors associated with other aspects of the production system influence the type and rate of disfluencies produced.

It also appears that lexical frequency may only have an influence on speakers' disfluency rates for very low frequency words. Disfluency rates for clear items in Experiment 2 were closer to those from the high frequency condition in Experiment 1 than the low frequency condition, which contained more than twice as many disfluencies. However, this could also be accounted for by differences in name agreement. Both of these factors could provide alternate accounts for the results of Experiment 1. If name agreement is the driving factor, disfluencies may be the result of increased processing requirements related to choosing one of several competing lexical representations for a particular item. Further research explicitly manipulating frequency and name agreement is necessary to resolve this question.

When disfluencies in the path description for Experiment 2 were examined, a clear effect of the number of paths was observed for all disfluency classes. However, the majority of disfluencies observed during description of the path were either fillers or repairs, suggesting that increases in processing load due to planning and formulating an utterance with multiple descriptive options may cause difficulties that often take some time to resolve. The increase in repairs suggests that often a speaker may embark upon planning and even generating an utterance before they know which path the marker is taking.

Although these studies highlight many issues for further research, they also provide a significant contribution to the ongoing debate concerning the provenance of disfluency. By explicitly manipulating the content of spontaneous speech, we provide causal evidence that particular disfluencies may be the consequence of different underlying difficulties in the language production system. Experimental approaches such as those illustrated here provide an important complement to the existing body of corpus-based disfluency literature.

Acknowledgements

The first author was supported by ESRC studentship PTA-030-2002-01229. We gratefully acknowledge the help of Lucy Clark, Hannah Furness and Lucy MacGregor.

References

Baayen, R. H., Piepenbrock, R., & van Rijn, H. (1993). The CELEX Lexical Database. Nijmegen: Centre for Lexical Information.

Barry, C., Morrison, C. M., & Ellis, A. W. (1997). Naming the Snodgrass and Vanderwart pictures: Effects of age of acquisition, frequency and name agreement. *Quarterly Journal of Experimental Psychology, 50A*(3), 560-585.

Bortfield, H., Leon, S., Bloom, J. E., Schober, M. F., & Brennan, S. E. (2001). Disfluency rates in conversation, age effects, relationship, topic, role and gender. *Language and Speech, 44*(2), 123-147.

Brennan, S. E., & Schober, M. F. (2001). How Listeners compensate for disfluencies in spontaneous speech. *Journal of Memory and Language, 44*, 274-296.

Clark, H. H., & Fox Tree, J. E. (2002). Using uh and um in spontaneous speaking. *Cognition, 84*, 73-111.

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

Griffin, Z. M., & Huitema, J. (1999). Beckman picture naming norms. From http://langprod.cogsci.uiuc.edu

Kilgariff, A. (1995). BNC database and word frequency lists. From http://www.itri.brighton.ac.uk

Kucera, H., & Francis, W. N. (1967). *Computational analysis of present-day American English*. Providence, RI: Brown University Press.

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

Maclay, H., & Osgood, C. E. (1959). Hesitation phenomena in spontaneous English speech. *Word, 15,* 19-44.

Oomen, C. C. E., & Postma, A. (2001). Effects of time pressure on mechanisms of speech production and self-monitoring. *Journal of Psycholinguistic Research, 30*(2), 163-184.

Rossion, B., & Portois, G. (2004). Revisiting Snodgrass and Vanderwart's object pictorial set: The role of surface detail in basic-level object recognition. *Perception, 33*, 217-236.

Schachter, S., Christenfeld, N., Ravina, B., & Bilous, F. (1991). Speech disfluency and the structure of knowledge. *Journal of Personality and Social Psychology, 60*, 362-367.

Siegman, A. W., & Pope, B. (1966). Ambiguity and verbal disfluencies in the TAT. *Journal of Consulting Psychology, 30*(3), 239-245.

Shriberg, E. (1996). *Disfluencies in Switchboard*. Paper presented at the International Conference on Spoken Language Processing (ICSLP '96), Philedephia, PA.

Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: norms for name agreement, image agreement, familiarity and visual complexity. *Journal of Experimental Psychology: Learning, Memory and Cognition, 6*, 174-215.

Swerts, M. (1998). Filled pauses as markers of discourse structure. *Journal of Pragmatics, 30*, 485-496.