

## Filled Pauses in Language Teaching: Why and How

Ralph L. Rose

### Abstract

Filled Pauses (*uh*, *um*) are ubiquitous elements of spontaneous speech but have received relatively little attention in second language teaching. Perhaps this is because filled pauses have often been regarded as meaningless elements resulting from speech processing difficulties. This paper draws from research in widely disparate fields to show that speakers and listeners use them systematically and meaningfully. These facts are used to generate a unified and coherent description of filled pauses in spontaneous speech. This is then used to develop a concept of communicative competence in which filled pauses play a role at the interface between pragmatic constraints and communication strategies. The article concludes with practical recommendations for how filled pauses may be incorporated into the second-language teaching curriculum.

One of the most common tokens in the Collins COBUILD corpus of spoken English language is forms of the first-person pronoun (Rose, 1998). Roughly two-thirds as common are forms of the third-person neuter pronoun, the indefinite article, and filled pauses (FPs; e.g., *uh*, *um*<sup>1</sup>). Presumably, all English language teaching texts give at least implicit attention to pronouns and articles, and any text which avoided third-person neuter pronouns or indefinite articles would arguably be of questionable value in the language classroom. However, surprisingly little attention has been paid to FPs in English language teaching materials. The reasons for this are unclear. There have been occasional calls in the literature for such attention (cf., Crystal and Davy, 1975; Di Pietro, 1980; Fayer and Krasinski, 1987; Leeson, 1970; Voss, 1979). There may be three reasons for the dearth of FPs in the language teaching curriculum. First, the common sense view of FPs is that they are used when the speaker is trying to decide what to say. In slightly more technical terms, they are a byproduct of speech planning. As such, they are considered devoid of communicative value—mere throwaway elements of spontaneous speech. While they may indicate speech production processes—and there is wide consensus on this (cf., Dalton and Hardcastle, 1977; Færch and Kasper, 1983; Goldman-Eisler, 1961)—it does not follow that they are meaningless. The mere fact that they indicate speech production processes is meaningful. Both speakers and hearers may exploit this fact when encoding or decoding speech, respectively. They may communicate or perceive something more than or different from what the surrounding words alone indicate. This kind of systematic use of FPs in spontaneous speech cannot be explained merely in terms of speech production or perception mechanisms, but must also be explained in terms of knowledge of the language.

A second reason why FPs have not been integrated into language teaching thus far may stem from the fact that the use of FPs in speech—particularly in public speaking situations—is stigmatized. It is widely believed that listeners form negative judgments about speakers who use FPs. The interlanguage of second language learners is already deficient—so the argument goes—so why should language teachers further handicap them with such overt elements of disfluency? Regardless of whether learners ever use FPs in their speech (and I will argue in this paper that they should be taught to do so to a limited extent), they will inevitably encounter native speakers who do use FPs regularly. In fact, it is likely that *every* native speaker they

---

<sup>1</sup>In the COBUILD Corpus, FPs are rendered as *er* and *erm*—the preferred orthographic rendering of these items in British English.

converse with at length will use FPs. Hence, it is useful to prepare the learner for understanding these forms.

However, even if FPs are to be incorporated in the language curriculum, teachers are still faced with the question of how to do so. There is little explicit guidance in the literature as to how FPs might fit into language instruction. This is the third reason which explains why there is little treatment of FPs in language teaching. Research on FPs (and related phenomena) has not been the exclusive domain of any particular field of language study. FPs have been investigated by clinical and behavioral psychologists, sociologists, theoretical and applied linguists, cognitive scientists, and speech-recognition technology researchers, forensic specialists among others. Furthermore, the literature spans some five decades. So, it is not surprising that this research has not been united into a single coherent presentation for those who are expected to present language in a pedagogically suitable manner.

This paper is designed to present an argument for *why* FPs should be a part of the second language teaching curriculum, and to propose *how* FPs may be incorporated. Toward these ends, a coherent description of FPs in communication is presented in which FPs are used by speakers as strategic devices—largely to compensate for linguistic processing difficulties—and that pragmatic constraints motivate the need to use them. I argue that the patterns of use of such devices are language-specific and that learners should therefore be taught to use FPs as they are used in the target language. As such, the teaching of FPs in the second-language curriculum constitute efforts to develop the pragmatic and strategic competence of learners.

## Attitudes Toward FPs: What is the Listener Attending to?

According to conventional wisdom, using FPs (particularly in public-speaking situations) reflects poorly on the speaker. A short talking paper on verbal fillers given at a local meeting of an international organization devoted to improving its members' public speaking skills had this to say about FPs: "Um' sounds dumb! 'Uh' sounds like 'duh!'" Even one ESL textbook (Viney and Viney, 1996) in a one-page treatment on "Thinking Time" makes the claim that "If you use [FPs] too often you sound stupid" (p. 79). While it is easy to find plenty of people who echo such claims, I can find no empirical evidence to support a negative correlation between FP use and perceptions of a speaker's intelligence.<sup>2</sup> However, there is some evidence which appears to correlate high FP use with certain negative judgments. Speakers with a high FP rate may be judged by listeners as less credible (McCroskey and Mehrley, 1969; Sereno and Hawkins, 1967), less truthful (Fox Tree, 2002; Kraut, 1978), less open (Fischer and Apostol, 1975), or less competent (Norton-Ford and Hogan, 1980). However, there are some limitations to the conclusiveness of these studies. Many of these studies have examined FPs as members of a larger class of hesitation phenomena (hereafter, HP: including such other items as false starts, restarts, and silent pauses; see Maclay and Osgood, 1959). HP have been shown not to be such a homogenous group. In particular, silent pauses show a very different pattern of occurrence than FPs. For instance, silent pauses (and other HP excluding FPs) are a much more reliable indicator of speaker anxiety than FPs (Goldman-Eisler, 1961; Kasl and Mahl, 1965; Krause and Pilisuk, 1961; Mahl, 1956; Ragsdale, 1976). It is conceivable, then, that listeners who form negative judgments of speakers with high hesitation rates are not necessarily basing this judgment on FPs, but on other HP, or some interaction between FPs and other HP.

Another problem with the studies cited above is that, in effect, they measured attitudes in public-speaking situations. While the recordings used may have been samples of spontaneous

---

<sup>2</sup>The only study I have been able to find which looks at hesitation and measures of intelligence is Bernstein (1962) in which no significant correlation was found between the two. On the other hand, social class was found to be a more significant factor where working-class youth were found to use less hesitation than middle-class youth. The reader is referred to Bernstein's paper for details of the study and the social theory of elaborated and restricted codes on which the study is based. Also see Coulthard (1969) for critique of the theory.

speech (in fact, many of them were not spontaneous but scripted—even the hesitations—a fact which raises doubts about authenticity), the experimental context, by broadcasting the recordings to an audience of non-interactive subjects, effectively made it a public-speaking event in which expectations about speech performance and interaction are very different. Furthermore, subjects knew or could reasonably infer that their task was to evaluate the speaker's speaking skills. This metacognitive task may require very different skills than those used during a lecture, or particularly than the skills used during conversational interaction. I can find no study that measures listeners' evaluational reactions in informal contexts.

Yet, it still is difficult to shake the intuition that FPs may have some sort of negative influence on a listener. In order to reconcile this, let's put ourselves in the role of a listener. It is easy to recognize that we do not always notice when speakers use FPs. In fact, it might even be the case that we rarely notice them (see Lickley, 1995, for experimental evidence of this). This phenomenon, sometimes described as *filtering* (Martin and Strange, 1968; Shriberg, 1994), may point toward an explanation. When we listen to a speaker, there are many things we may attend to: We may notice the individual words, the particular syntactic structures, the pronunciation of individual phonemes, the prosodic structure, the message content, or the discourse organization. We may even notice such non-linguistic things as gestures, body position, or even what color hair the speaker has. However, much of the time, we are primarily focused on the message that the speaker intends to communicate, while the other information is processed only to the degree that it is necessary to determine message content and then is disposed of. Yet, if the speaker chooses a particularly obscure word, or commits a relatively egregious grammatical error, or even has a bad hair day, then we might attend to these more and perhaps judge the message or the speaker differently.

In the studies cited above which purport to show negative judgments of speakers who use FPs, the subjects' attention was not controlled. Hence, it is not clear just why they gave negative reactions: Is it that FPs cause an unconscious negative reaction in listeners, or is it that listeners attending to FPs in speech give negative reactions? Christenfeld (1995) examined this question by carefully controlling what listeners were asked to attend to in speech. Subjects were instructed to listen to either content or style (of presentation) in three pause conditions: with FPs, with silent pauses, and with no pauses. Christenfeld observed that when attending to the content, subjects judged that speech with no pauses was most eloquent. However, interestingly, there was no difference in judgments of eloquence between speech with FPs and speech with silent pauses. Furthermore, there was no difference in the judgments of relaxedness between speech with FPs and speech with no pauses. Speech with silent pauses was judged significantly less relaxed than the other two. On the other hand, when subjects attended to style, eloquence judgments were lower overall. However, speech with FPs was judged more relaxed than the other two. In short, speech with no pauses seems to be the best, but interestingly, there is no advantage to replacing FPs with silent pauses in speech. In fact, using silent pauses may sound *more* anxious (consistent with the studies cited above).

Returning to the original question of whether FPs reflect poorly on the speaker, the answer seems to be a qualified no: it is not *inherently* bad to use FPs. Doing so does not guarantee a negative reaction in listeners. However, it appears that there may be negative consequences when listeners *notice* FPs. So, the real challenge for a speaker is how to keep the listener focused on the content of the talk and not the style. (Christenfeld, 1995) concludes

When an audience attends to style, it may well be the result of the content being unworthy of attention, or the speaker's style being distracting. In this case, ums will not be associated with poor speech, but noticing ums will be. Just about every speaker produces ums, but the good speakers, by keeping substance, not style, as the center of attention, will effectively hide their hesitations. (p. 185)

Second language learners—given that they are at varying levels of proficiency in the target

language—will need to hesitate in some way or another when they speak. Instructing students to just ‘speak fluently’ is not helpful here because it fails to address the problem: They are having a production problem and need some means to deal with it. In addition, instructing them to use silent pauses is also insufficient for reasons noted above (and because of the risk of interruptions, discussed below). Encouraging them to use FPs *judiciously* then becomes the next best choice.

Finally, one more comment on this topic is necessary. Using FPs in a public-speaking situation—for which one has presumably had time to prepare—may create the environment for negative speaker evaluation. But in conversational interaction, the consequences of FPs are unknown. If the notion of filtering bears much weight here, then it would seem that FPs in spontaneous speech in a conversational context are less likely to be noticed and hence do not carry the same sort of implications as do FPs in more formal contexts.

## Crosslinguistic Variation in FP Pronunciation

Even if one is still not persuaded that the line of reasoning given above is sufficient motivation for teaching FP use in the language classroom, the research described above does suggest at least a minimal treatment of FPs as follows. FPs exhibit surprising regularity in patterns of pronunciation within a language. In English, the predominant patterns use a monosyllabic structure with the mid-central vowel: /ə/ or /əm/. French speakers use an open mid-front rounded vowel, /œm/, while Japanese speakers use a bisyllabic construction, /ε-to/ (see Clark and Fox Tree, 2002, for other examples). A little reflection shows that these patterns are regularized and that variation from them sounds odd. Consider a FP in English pronounced with a low-back vowel as in /am/, or even with a high-front vowel, /im/. These variants are much more noticeable than the conventional pronunciations in the same way that dialectal variation in the pronunciation of, say, *girl*, are easily noticeable. However, the experimental observations of Christenfeld (1995) suggest the risk that listeners who hear these non-conventional variants will then begin to focus on the speaker’s style and will be distracted from the content. In Fayer and Krasinski (1987), native English speaker reactions to hearing the English speech of native speakers of Spanish were observed. The listeners indicated HP, along with pronunciation, to be the most distracting from the message. An explanation for this could be that HP pattern discrepancies from target language norms caused listeners to notice the discrepancies and be distracted from the message. So, the logical conclusion seems to be that learners who feel the urge to use FPs in their speech—and at some point or other, all learners will feel this urge—should be encouraged to use target language patterns of FP pronunciation.

## FPs in Interaction

In the previous sections, I have tried to dispel, or at least clarify, some myths about listeners’ evaluational reactions to FPs in spontaneous speech. In this section, I examine the communicative value of FPs in interaction. While it is relatively uncontroversial that FPs are indicative of speech production processes, this is a psychological explanation and is only part of the story. Language production occurs in a social interactional context which imposes certain pragmatic constraints on speakers. Here, I will make the case that such constraints motivate the use of FPs in spontaneous speech and further that cultural interactional norms require speakers to exhibit this hesitation in certain contexts.

## Mitigating Devices

In the following pair of interchanges, there is a distinct difference in the tone of the reply given by speaker B.

- (1) A: Would you like to go to the movies?  
B: No, thanks.
- (2) A: Would you like to go to the movies?  
B: Uh ... no, thanks.

B's response to A's invitation in (1) appears quite sharp and abrupt. The form of the declination strongly suggests that B dislikes A or doesn't wish to be with A. In contrast, the effect of the declination is softened in (2). The FP here almost seems to say, "Please get ready, I'm about to decline your invitation" (cf., Davidson, 1985). This contrast can be explained in terms of the preservation of harmony Brown and Levinson (1987); Goffman (1967, cf., face and politeness theory in). In conversational interaction, interlocutors are constrained by social norms to preserve harmony between them. A harmonious response—and therefore a preferred response—to an invitation is an acceptance. As such, an economical acceptance form is sufficient: "Sure!" However, when the invitee must decline, additional effort must be made to preserve harmony. This additional effort triggers an increased cognitive load as the invitee plans an appropriate harmony-preserving response—perhaps by giving an account of why the invitation must be declined. If these processes are too burdensome, then hesitation may result, often in the form of FPs.

Mitigating devices occur in a wide range of contexts. For instance, Eakins and Eakins (1978) observed that FPs are more frequent when speakers are trying to downplay their assertiveness. An investigation by Norton-Ford and Hogan (1980) further showed that listeners apparently detect these signals and make corresponding judgments of (non)assertiveness or (non)aggressiveness. Brennan and Williams (1995) observed that speakers often used mitigating devices (typically FPs) when responding to factual questions (e.g., "What is the capital city of Australia?") with non-answers (e.g., "I don't know."). They further observed that listeners to these responses judged that speakers of mitigated (i.e., hesitant) non-answers were more likely to have known the answer to the question (but temporarily forgotten it) than those who gave unmitigated non-answers. In other words, when somebody said, "Um, I don't know", listeners were more apt to interpret the FP as an indication that they knew the answer but had some difficulty retrieving it from memory.

Another use of FPs as mitigating devices deserves comment. The spontaneous nature of conversation means that hesitations and errors will occur. Clark (1994) argues that speakers manage these by giving an overt account of the delay in their speech. In other words, FPs may serve as a warning to interlocutors that the speaker needs a little more time to formulate the current utterance. Because a speaker has control of the conversation, he or she has an obligation to continue speaking until the end of a turn. This mitigating use of FPs may serve as a sort of plea for patience or even for assistance. Interestingly, research shows that listeners are sensitive to the FPs used in the middle of such error sequences. In such an utterance as, "Move it to the yellow uh purple square", the FP appears to be an overt pragmatic cue—a mitigating device—that the speaker is preparing to utter a repair. The listener may take strategic advantage of this by being prepared for such a repair. Contrast this case with that in which a speaker utters, "Move it to the yellow purple square." Here there is no overt warning of a repair, only the repair itself—which the listener must recognize and interpret as such. This requires more effort by the listener than when a FP is used. This was demonstrated by Brennan and Schober (2001) in a psycholinguistic investigation. Subjects found the correct target faster when there was an overt repair warning—that is, a FP—than when there was none, or even than when there was no repair at all (i.e., "Move it to the purple square."). Listeners have developed a certain sensitivity to FPs in their speech perception to deal with and actually take advantage of these mitigating devices in spontaneous speech.

The interaction between the pragmatic constraints to preserve social harmony with the cognitive processes underlying speech production lead to a systematic pattern of FP use as mitigating devices. While the preservation of social harmony may be universal, the precise set

of pragmatic constraints that achieve this end may vary across languages. A good illustration is the cultural valuation of silence—a topic closely related to FPs. Many sociolinguists have noted widely differing cultural values on silence (cf., Basso, 1972; Tannen, 1985). In some cultures, even a little silence causes discomfort, while in other cultures, long stretches of silence are not only comfortable, in some cases they are obligatory. Language learners need training to realize how such pragmatic constraints lead to context-appropriate hesitations.

## Discourse Management

Conversation analysts have long noted that FPs have a high rate of occurrence in two prominent locations in discourse structure: at discourse segment boundaries, and at the beginning of conversational turns. In this section I examine how FPs serve to indicate discourse structure and how this fact satisfies certain pragmatic constraints on cooperation during interaction.

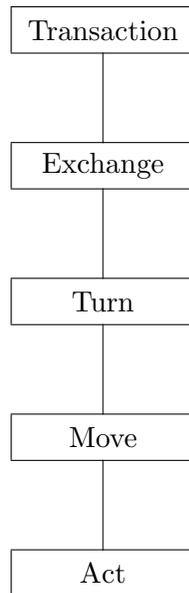
## Discourse Boundaries

Discourse structure is typically conceptualized as a hierarchical structure in which each level of the hierarchy may contain one or more instances of the level below it on the hierarchy. For instance, Stenstrom (1994), adapting from Sinclair and Coulthard (1975), defines spoken discourse structure in terms of five hierarchically arranged levels of discourse segments (see Figure 1)—*transaction*, *exchange*, *turn*, *move*, and *act*—where each level contains one or more instances of the segment immediately below it on the hierarchy. The cognitive representation of each level comprises a set of relevant properties including the discourse purpose of that level and some sort of map of the ordering of the levels it immediately contains. In speech, before crossing any particular discourse boundary—regardless of the level—the speaker must do some planning of that discourse segment. Furthermore, the planning of any discourse segment entails planning of all hierarchically lower segments it contains. Hence, beginning a new turn, for example, entails much greater planning than beginning a new act. So, the prediction is that higher (i.e., levels higher on the hierarchy) discourse boundaries should exhibit signs of greater language planning. Thus, we would expect to see more frequent occurrence of FPs at, say, an exchange boundary than at an act boundary. This prediction was born out in Swerts (1998) who compared the occurrence of FPs at “strong” and “weak” discourse boundaries. He found that the initial prosodic phrase following a strong discourse boundary was more likely to have an initial FP than either a medial FP or no FP at all. On the other hand, the phrase following a weak discourse boundary was least likely to have an initial FP.

In a psycholinguistic study of the perception of short discourses, Bailey and Ferreira (2003) observed that listeners are sensitive to these phenomena. Using such sentences as *Sandra bumped into the busboy and the uh uh waiter told her to be careful* and *Sandra bumped into the busboy and the waiter uh uh told her to be careful*, they found that when the FPs came before the head noun (i.e., *the waiter*) of the second clause, subjects preferred to interpret the noun as the beginning of a new clause rather than as (part of) the direct object of the preceding clause. So, in the linear process of language perception, if an FP occurs at a spot which is conceivably a discourse boundary, then listeners prefer to consider it a discourse boundary.

One of the fundamental principles of pragmatics is the cooperative principle of Grice (1975): “Make your conversational contribution such as is required, at the stage at which it occurs, by the accepted purpose or direction of the talk exchange in which you are engaged” (p. 67). Grice distinguishes four sub-categories of this principle. I would like to focus briefly on one, the maxim of *manner* which states, among other things, “Be orderly” (p. 67). In marking discourse boundaries, FPs support this maxim, by cooperatively informing the listener how the sequence of utterances being spoken should be organized into meaningful discourse units. This function may of course be accomplished by such organizational devices as *first*, *next*, *then*, *last*, or such

Figure 1: Discourse Structure Hierarchy



connectives as *therefore* or *on the other hand*. Speakers also use FPs in a systematic fashion to mark discourse structure cooperatively, this is systematically understood by listeners.

### Conversational Turns

Another area of discourse organization in which the cooperative principle motivates the use of FPs is in the management of conversational turns (Finegan, 1994; Levinson, 1983; Sacks et al., 1974; Stenstrom, 1994). A FP may be used in order to establish control of the conversational “ball” (borrowing a metaphor from Maclay and Osgood, 1959). It may even be used to wrest control of the conversation away from another (Murata, 1994). A FP can be used to “hold” a conversational turn—to prevent anyone else from taking control: If a speaker “pauses long enough to receive the cue of his own silence, he will produce some kind of signal ([m, er]...) which says, in effect, ‘I’m still in control—don’t interrupt me!’” (Maclay and Osgood, 1959, p. 41). Two separate investigations (Ball, 1975; d’Urso and Zammuner, 1990) demonstrated this in elicited conversations, finding that an utterance followed by a FP and then by silence prevented a listener from assuming a turn longer than an utterance followed by mere silence. Beattie (1977) made similar observations in analyses of naturally-occurring conversation: Interruptions were significantly more likely during silent pauses than during FPs.

In the same way that speakers are obliged to outline the discourse structure of their speech, they must also cooperatively signal to others the beginning and end of their conversational turns. In public speaking situations, where the speaker has been ceded the right to speak at length without interruption, silent pauses may be longer or more frequent. However, in interactive situations, interlocutors are less constrained and may interrupt more freely. Then it is important for the speaker to proactively, yet cooperatively, maintain control of the conversation. FPs accomplish this task.

### Discourse Status of Entities

While FPs at discourse boundaries and in conversational turn management have been discussed in the literature for some time, there is one recent discovery from the psycholinguistic literature about the role that FPs play in discourse. Given that FPs coincide with increased cognitive

effort in spontaneous speech production, what can we conclude about FPs that occur with a noun phrase (e.g., *uh the cat* or *the uh cat*)? Apparently, additional cognitive effort is being spent to recover the name of the noun from memory. That suggests that it is unlikely that the entity is currently stored in short-term memory where access should be easy. Rather, it must be recovered from long-term memory. In terms of the given-new distinction in discourse, then, this means the entity must be a new discourse entity (or so old that it has recessed from short-term memory). Arnold et al. (2004) observed subjects making these judgments about the discourse-status of entities. Subjects preferred to interpret a noun phrase with a FP (e.g., *the uh cat*) as referring to a discourse-new entity; Apparently they interpreted the FP as an indication that the subject was searching long-term memory for the name of the entity. Here, again, we see how both speakers and listeners use FPs cooperatively to manage the communication of a discourse.

In recent years, there have been increasing calls for instruction in pragmatics in language curricula (c.f., Bardovi-Harlig, 2001). The evidence here shows that the pragmatic concept of the cooperational principle explains how FPs are used by speakers and listeners in the management of discourse during spontaneous speech. As such, instruction on FPs should be useful in the syllabus.

## FPs and Communicative Competence

Above, I have argued that FPs in spontaneous speech are used by both speakers and listeners in ways that are systematic and meaningful. Instead of being mere interruptions in message transfer, they are symbolic elements of communication. In particular, their use is motivated by the application of various pragmatic constraints which result in increased cognitive effort. Knowledge about the appropriate and inappropriate use of FPs in speech must constitute some part of a speaker's communicative competence. This section describes how FPs contribute to overall communicative competence.

Canale and Swain (1980) characterize communicative competence as consisting of three components: *grammatical*, *sociolinguistic*, and *strategic* competence. Grammatical competence comprises mastery of the linguistic code. Sociolinguistic competence consists of knowledge of the appropriate use of language in a given setting. Finally strategic competence comprises the skills that help overcome shortcomings in grammatical competence. In the exercise of their strategic competence, speakers draw from an inventory of communication strategies—a set of devices which help them manage problems or even avoid problems in communication. This may include such linguistic techniques as circumlocution or even non-linguistic techniques as gesture.

One class of communication strategies which has been described in the literature is fillers and hesitation devices (Dörnyei, 1995; Dula, 2001). These strategies compensate for speech processing problems by allowing the speaker to fill or gain time during speech and thereby fulfill the goal of making a complete contribution by giving some vocal indication that a continuation is forthcoming. I argue that FPs fit into this class and fulfill the same function as such lexical fillers and hesitation devices as *well*, *hmm*, *let's see*, *a-nd*, and *so-*.

Dörnyei (1995) and Dula (2001) give persuasive evidence that fillers and hesitations can be taught as communication strategies with positive results. However, they apparently restrict their inventory of fillers and hesitation devices to lexical ones.<sup>3</sup> This seems an overly restrictive definition as FPs can occur in every environment in which a lexical hesitation device can occur. But this fact is not reciprocal. Consider (3)-(4).

(3) A: Excuse me, what time is it?

---

<sup>3</sup>Both Dörnyei (1995) and Dula (2001) make only brief reference to FPs and do not clarify whether or not they were actually measured in their respective investigations.

- B: Um / Let's see, it's 2:45.  
(4) A: Where are you going?  
B: I'm going to the uh store.  
B': ?I'm going to the let's see the store.

The B sentence in (3) is perfectly natural beginning with either a lexical hesitation device or a FP. However, (4)B' is marked (indicated typographically with a question mark) and if it does mean something, it does not mean the same thing as (4)B. So, if anything, it may be more correct to regard these lexical hesitation devices as members of the class of FPs!<sup>4</sup> While that question is interesting and may have certain implications for how verbal and nonverbal phenomena interact, I reserve that for a later paper. What I am advocating is that FPs along with fillers and hesitation devices form a functionally coherent subcategory of communication strategies.

Furthermore, I maintain that although FPs are used to achieve hesitation, this is not a complete characterization of how FPs contribute to overall communicative competence. There may be pragmatic constraints that cause the speaker to need more processing time. Knowledge of these constraints and how they interact with communication strategies to produce FPs is a part of a speaker's sociolinguistic competence. In particular, knowledge of interactional norms of politeness and discourse also influence the use of FPs in conversation.

In short, FPs are meaningful in the exercise of sociolinguistic and strategic competence in spontaneous speech.

## FPs in the Language Classroom

Now we consider the practical question of how to treat FPs in the second-language classroom. First, I would like to outline four general principles about such treatment.

### Four Principles

#### **Use target language pronunciation of FPs.**

Learners should be encouraged to draw from the target language inventory of FP pronunciations. The use of native-language FPs in target language speech may cause listeners to notice the imported FP, thereby distracting them from the content. It may even be useful to present FPs as vocabulary items. Although it remains an open question whether FPs are words in English (see Clark and Fox Tree, 2002, for a persuasive argument that they are), in other languages, the most common FPs are certainly more word-like, often similar to (or possibly derived from) demonstratives (e.g., Japanese *ano-* and Spanish *este*). For speakers of these languages, it may be easier to conceptualize English FPs as words and to learn them as such with their conventionalized pronunciations.

#### **Do not rely on native language transfer.**

Some may now claim that although the role of FPs in a speaker's communicative competence may have value, this doesn't necessarily mean that we need to teach FPs in the classroom. All languages have FPs and the cognitive processes that govern the use of FPs will transfer readily to the target language. This is analogous to the argument sometimes used to question the usefulness of teaching communication strategies at all (cf., Kellerman, 1991). However, that reasoning does not take into account the possibility of cross-linguistic variation in the appropriateness of these strategies with respect to pragmatic constraints. Silence, as discussed

---

<sup>4</sup>This is the approach taken in Rose (1998) in which FPs were broadly classified as vocalized hesitations and subclassified as either lexicalized or unlexicalized.

above, provides an example in which the interaction between cultural values and the kinds of strategies available to fill silence is language-specific and may not transfer automatically. Furthermore, the actual inventory of possible FPs clearly does vary cross-linguistically and does not transfer any more than a French learner of English can transfer the native *chapeau* to the target *hat*. However, since the availability of hesitation strategies appears to be pervasive cross-linguistically, this may be a useful point of reference for teaching about FPs in the second-language curriculum.

### **Give early, varied, and frequent Exposure**

FPs can provide significant perceptual difficulties for nonnative listeners. In an experiment involving transcription of target language speech, Voss (1979) observed that nonnative listeners may confuse FPs and the indefinite article because of their homophonous nature (i.e., *uh* as *a*, *um* as *an*). Alternatively, FPs may be misunderstood as parts of adjacent words as in (5).

- (5) spoken text (by native):                    "A contract, uh, is when ... "  
transcribed text (by nonnative):        "A contractor is when ... "

(author's own data)

Therefore it is crucial to expose the language learner to naturally-occurring FPs in listening materials or in printed scripts intended to represent spontaneous speech. The exposure should begin early. There is no reason why beginners should not begin to be exposed to FPs immediately.<sup>5</sup> Furthermore, this exposure should exhibit FPs in a wide variety of social and linguistic contexts. The few ESL textbooks which do illustrate FPs in spontaneous speech tend to include them only as discourse boundary markers, or as mitigating devices in invitation refusals (although the mitigating function is rarely noted). FPs should be presented in the many contexts in which they are found—phrase-medial and phrase-final as well as phrase-initial. Finally, FPs should be presented at frequency levels close to that of native speech. Nine FPs per minute has been called a "normal" rate (Christenfeld, 1995). However, it need not be the case that all materials have FPs at this rate. Certainly, pedagogical reasons may dictate a lower rate in certain materials. Nevertheless, at some point, learners need to be prepared to hear FPs as they will hear them in the actual discourse of native speakers.

### **Show FP variation across speaking contexts**

The research on evaluational reactions to FPs in speech suggest that listeners in formal speaking contexts are more likely to notice FPs and make negative evaluations of speakers than those in informal, conversational interactions. It is useful then to consider the anticipated speaking needs of students in choosing an approach to FPs in the language classroom. Those who are studying business English, or practicing delivery of a speech at a conference may require different guidance than those who are preparing to travel abroad or join a homestay program. The treatment of FPs should be tailored to respond to the varying needs and goals of the learners.

### **Suggested Approaches and Activities**

In this section I outline some specific approaches and activities that may be taken in the classroom or in language teaching materials for teaching FPs.

---

<sup>5</sup>Rose (1998) argued against teaching FPs to beginners because of the possibility of perceptual difficulties. However, I now believe this is exactly the reason why beginners should be exposed to FPs as soon as possible: in order to reduce subsequent problems.

## Raising Awareness of FPs in Target Language Speech

While exposing learners to FPs as described above should give them greater awareness of FPs and their role in spontaneous speech, it may be useful to incorporate explicit practice designed to help listeners to filter FPs in favor of the central message. Next are two activities designed to develop these filtering strategies.

**Recognition exercise** An early step towards developing learners' filtering ability is to help them recognize the presence of FPs in native speech. A recognition exercise can accomplish this. Learners listen to a spoken text containing FPs while following a printed transcript of the text without the FPs. Learners are then asked to mark the locations of FPs in the transcript.

**Paraphrase** Once learners begin to recognize FPs, they need to practice filtering them. This can be accomplished by having students listen to a recording of spontaneous speech with FPs and asking them to express the essential meaning of the passage, removing FPs.<sup>6</sup> For example, learners might listen to the following spoken text.

- (6) Okay this is um a man and a woman um at home. And uh the man falls down the stairs and he hurts himself. Maybe he broke his leg or something. And uh the woman is very concerned tries to help him up but she can't she can't uh he can't get up you know. He's uh his leg hurts too much. So so she uh she calls probably uh an ambulance or something. And tells him not to worry that it'll be okay.

(author's own data)

Learners would then be expected to produce the following in spoken or written form.

- (7) This is a man and a woman at home. The man falls down the stairs and hurts himself ...

## Stalling

Many useful exercises can be described to help learners to develop stalling skills. A typical exercise might be to use a quiz show format in which the teacher asks each student a challenging question in turn and the student is required to give a fluent answer (where fluency here is defined as containing no long silent pauses). Students may use devices as shown in Table 1 to achieve fluency (see Rose, 1999, for a detailed description of this activity).

Table 1: Expressions Used to Gain Thinking Time

Stalling expressions	Asking for repetition	Concession
uh	Excuse me, could you say that again please?	I'm not sure.
um	Would you say that again?	I don't know.
hmm	What did you say?	
well	Excuse me?	
let me see	Pardon?	
	What?	
	Huh?	

<sup>6</sup>This activity was inspired by a similar activity in Dörnyei and Thurrell (1991). See their paper for further exercises which may readily be adapted for FP training.

## Mitigating Devices

There are a wide variety of social situations in which the need for mitigating devices is high: apologizing, suggesting, inviting, stating an opinion, disagreeing, criticizing, and making a complaint, to name a few. Instructional materials can raise learner awareness of the sociolinguistic use of FPs as mitigating devices by incorporating them into lists of possible expressions for each situation. For instance, an activity which covers making, accepting, and declining invitations might include a list of useful expressions as shown in Table 2.

Table 2: Useful Expressions when Making, Accepting, or Declining an Invitation

Inviting	Accepting	Declining
I was wondering if you would like to...	Sure, sounds great!	Uh...I'm sorry, I can't.
Would you like to...?	Yes, I'd like to.	Well, I'd like to, but...
How about...?		

With some students, an inductive approach to the study of invitations may prove effective. After presenting a variety of authentic invitation exchanges, learners can be asked to outline the differences between the acceptances and the declinations. One likely observation is that while acceptances tend to be short and immediate, declinations are longer and more hesitant, and furthermore that this hesitation is often accomplished with FPs.

## Conversational Turn-Taking

An important skill for a language learner to master is that of turn-taking in interaction. It is also important to realize that these strategic skills are constrained by sociolinguistic norms of interaction. For instance, it is possible to take control of a conversation by raising one's voice, or by using taboo language. However, depending on the social context, these strategies may have widely varying effects, some not harmonious. As discussed above, FPs provide a widely useful strategy which conforms to sociolinguistic norms in conversational turn-taking. A list of a variety of such strategies incorporating FPs is shown in Table 3.

Table 3: Useful Expressions for Discourse Management

Taking a turn / Interrupting	Holding a turn	Ending a turn	Changing the topic
Uh/Um...	And...	<silence>	Uh/um...
Hmm...	So...	What do you think, Mary?	By the way...
Well...	Uh/Um...	Did you hear about...?	

One highly motivational exercise of turn-taking skills incorporates an element of competition. After agreeing on a topic, pairs engage in a five-minute discussion. While one student is speaking, the other may not interrupt or take a turn unless there is undue silence or a nomination (i.e., a forced change of turn by, say, asking a question) from the current speaker. The winner is the person who maintains the control of the conversation for a longer total time.

## Evaluation

Finally, it is useful to consider how research into the nature of FPs in spontaneous speech is relevant to how a language learners' communicative competence is evaluated. Ultimately, the standard must be the communicative competence of the native speaker. Research shows that native speakers have highly-developed sociolinguistic skills in which hesitation in the form of

FPs can be important. Furthermore, native speakers have a variety of strategies available to them in dealing with the time demands of either speech production or perception. These are the standards by which we must judge the communicative competence of English learners: to what degree do they exercise these same sociolinguistic skills and strategies as native speakers do? Incorporation of these facts into notions of fluency is also warranted. Fluency may be viewed not as message transfer without breaks or hesitation, but rather message transfer with context-appropriate hesitation (see Guillot, 1999, for extended discussion of how hesitation contributes to fluency).

## Conclusion

Filled pauses occur widely in spontaneous speech. However, rather than being annoying intrusions in speech, FPs are used by interlocutors in systematic and meaningful ways. Language teachers should prepare students to use and perceive FPs appropriately as do native speakers. I have presented a coherent description for how this may be done in the language teaching curriculum. However, I do not claim the approach presented is exhaustive and I hope others will add to this work. However, even more immediate is the need to try these ideas in practice to see how effective they are at improving learners' hesitation strategies. But how should this effectiveness be measured? It should not be considered successful if after a course of intensive study, the only measured result is that learners use more FPs in their speech. This is a weakness in the design of some earlier studies on the teachability of filler and hesitation devices as communication strategies (Dörnyei, 1995; Dula, 2001). While these studies have accomplished much, they do not go quite far enough. An effective case might demonstrate that learners use more FPs and other hesitation devices. More importantly though, an effective case should also demonstrate that native speakers perceive an improvement in certain aspects of the learners' communicative competence—perhaps most notably, fluency. That would constitute a successful case in the teaching of the production of FPs. In contrast, a successful case in the teaching of the perception of FPs might show that learners have less perceptual difficulty resulting from FPs in speech, and that they show the same sort of pragmatic sensitivity to the occurrence of FPs in speech.

Successful language teaching helps learners to speak and understand the target language as native speakers do. The production and perception exercises I outline above are designed to accomplish this goal by building on a coherent and justified understanding of how FPs are used in everyday speech.

## References

- Arnold, J., Tanenhaus, M., Altmann, R., and Fagnano, M. (2004). The old and thee, uh, new: Disfluency and reference resolution. *Psychological Science*, 15:578–582.
- Bailey, K. and Ferreira, F. (2003). Disfluencies affect the parsing of garden-path sentences. *Journal of Memory and Language*, 49:183–200.
- Ball, P. (1975). Listener's responses to filled pauses in relation to floor apportionment. *British Journal of Social & Clinical Psychology*, 14:423–424.
- Bardovi-Harlig, K. (2001). Evaluating the empirical evidence: Grounds for instruction in pragmatics? In Rose, K. and Kasper, G., editors, *Pragmatics in Language Teaching*, pages 13–32. Cambridge University Press, Cambridge, UK.
- Basso, K. (1972). "To give up on words": Silence in Western Apache culture. In Giglioli, P., editor, *Language and Social Context*, pages 67–86. Penguin, Harmondsworth, UK.

- Beattie, G. (1977). The dynamics of interruption and the filled pause. *British Journal of Social & Clinical Psychology*, 16:283–284.
- Bernstein, B. (1962). Linguistic codes, hesitation phenomena and intelligence. *Language and Speech*, 5:31–46.
- Brennan, S. and Schober, M. (2001). How listeners compensate for disfluencies in spontaneous speech. *Journal of Memory and Language*, 44:274–296.
- Brennan, S. and Williams, M. (1995). The feeling of another's knowing: Prosody and filled pauses as cues to listeners about the metacognitive states of speakers. *Journal of Memory and Language*, 34:383–398.
- Brown, P. and Levinson, S. (1987). *Politeness: Some Universals in Language*. Cambridge University Press, Cambridge.
- Canale, M. and Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics*, 1:1–47.
- Christenfeld, N. (1995). Does it hurt to say um? *Journal of Nonverbal Behavior*, 19:171–186.
- Clark, H. (1994). Managing problems in speaking. *Speech Communication*, 15:243–250.
- Clark, H. and Fox Tree, J. (2002). Using *uh* and *um* in spontaneous speaking. *Cognition*, 84:73–111.
- Coulthard, M. (1969). A discussion of restricted and elaborated codes. *Educational Review*, 22:38–50.
- Crystal, D. and Davy, D. (1975). *Advanced Conversational English*. Longman, London.
- Dalton, P. and Hardcastle, W. (1977). *Disorders of fluency*. Edward Arnold, London.
- Davidson, J. (1985). Subsequent versions of invitations, offers, requests, and proposals, dealing with potential or actual rejection. In Atkinson, J. and Heritage, J., editors, *Structures of Social Action: Studies in Conversation Analysis*, pages 102–108. Cambridge University Press, Cambridge, UK.
- Dörnyei, Z. (1995). On the teachability of communication strategies. *TESOL Quarterly*, 29:55–85.
- Dörnyei, Z. and Thurrell, S. (1991). Strategic competence and how to teach it. *ELT Journal*, 45:16–23.
- Di Pietro, R. (1980). Verbal strategies: A neglected dimension in language acquisition studies. In Dechert, D. and Raupach, M., editors, *Temporal Variables in Speech: Studies in Honour of Frieda Goldman-Eisler*. Mouton, The Hague.
- Dula, E. (2001). *The Effects of Communication Strategy Training on Foreign Language Learners at the University Level*. PhD thesis, Temple University.
- d'Urso, V. and Zammuner, V. (1990). The perception of pause in question-answer pairs. *Bulletin of the Psychonomic Society*, 28:41–43.
- Eakins, B. and Eakins, R. (1978). *Sex Differences in Human Communication*. Houghton Mifflin, Boston.

- Færch, C. and Kasper, G. (1983). On identifying communication strategies in interlanguage production. In Færch, C. and Kasper, G., editors, *Strategies in Interlanguage Communication*, pages 210–238. Longman, London.
- Fayer, J. and Krasinski, E. (1987). Native and nonnative judgments of intelligibility and irritation. *Language Learning*, 37:313–326.
- Finegan, E. (1994). *Language: Its structure and use*. Harcourt Brace, New York, second edition.
- Fischer, M. and Apostol, R. (1975). Selected vocal cues and counselor's perceptions of genuineness, self-disclosure, and anxiety. *Journal of Counseling Psychology*, 22:92–96.
- Fox Tree, J. (2002). Interpreting pauses and ums at turn exchanges. *Discourse Processes*, 34:37–55.
- Goffman, E. (1967). *Interaction Ritual*. Penguin, Harmondsworth, UK.
- Goldman-Eisler, F. (1961). A comparative study of two hesitation phenomena. *Language and Speech*, 4:18–26.
- Grice, H. (1975). Logic and conversation. In Cole, P. and Morgan, J., editors, *Syntax and Semantics 3: Speech Acts*. Academic Press, New York.
- Guillot, M.-N. (1999). *Fluency and Its Teaching*. Multilingual Matters, Clevedon, England.
- Kasl, S. and Mahl, G. (1965). The relationship of disturbances and hesitations in spontaneous speech to anxiety. *Journal of Personality and Social Psychology*, 1:425–433.
- Kellerman, R. (1991). Compensatory strategies in second language research: A critique, a revision, and some (non)-implications for the classroom. In Phillipson, R., Kellerman, E., Selinker, L., Sharwood-Smith, M., and Swain, M., editors, *Foreign/Second Language Pedagogy Research: A Commemorative Volume for Claus Færch*, pages 142–161. Multilingual Matters, Clevedon, England.
- Krause, M. and Pilisuk, M. (1961). Anxiety in verbal behavior: A validation study. *Journal of Consulting Psychology*, 25:414–419.
- Kraut, R. (1978). Verbal and nonverbal cues in the perception of lying. *Journal of Personality & Social Psychology*, 36:380–391.
- Leeson, R. (1970). The exploitation of pauses and hesitation phenonema in second language teaching: Some possible lines of exploration. *Audiovisual Language Journal*, 8:19–22.
- Levinson, S. (1983). *Pragmatics*. Cambridge University Press, Cambridge, UK.
- Lickley, R. (1995). Missing disfluencies. In *Proceedings of International Congress of Phonetic Science*, volume 4, pages 192–195, Stockholm.
- Maclay, H. and Osgood, C. (1959). Hesitation phenomena in spontaneous English speech. *Word*, 15:19–44.
- Mahl, G. (1956). Disturbances and silences in the patient's speech in psychotherapy. *Journal of Abnormal and Social Psychology*, 53:1–15.
- Martin, J. and Strange, W. (1968). The perception of hesitation in spontaneous speech. *Psychonomic Journals: Perception & Psychophysics*, 3:427–438.
- McCroskey, J. and Mehrley, R. (1969). The effects of disorganization and nonfluency on attitude change and source credibility. *Speech Monographs*, 36:13–21.

- Murata, K. (1994). Intrusive or co-operative? A cross-cultural study of interruption. *Journal of Pragmatics*, 21:385–400.
- Norton-Ford, J. and Hogan, D. (1980). Role of nonverbal behaviors in social judgments of peers' assertiveness. *Psychological Reports*, 46:1085–1086.
- Ragsdale, J. D. (1976). Relationship between hesitation phenomena, anxiety and self-control in a normal communication situation. *Language and Speech*, 19:257–265.
- Rose, R. (1998). The communicative value of filled pauses in spontaneous speech. Master's thesis, University of Birmingham, Birmingham, UK.
- Rose, R. (1999). Silence is golden? *TESOL Journal*, 8:35.
- Sacks, H., Schegloff, E., and Jefferson, G. (1974). A simplest systematics for the organization of turn-taking in conversation. *Language*, 50:696–735.
- Sereno, K. and Hawkins, G. (1967). The effects of variations in speakers' nonfluency upon audience ratings of attitude toward the speech topic and speakers' credibility. *Speech Monographs*, 36:59–64.
- Shriberg, E. (1994). *Preliminaries to a theory of speech disfluencies*. PhD thesis, University of California, Berkeley.
- Sinclair, J. and Coulthard, M. (1975). *Toward an analysis of discourse*. Oxford University Press, London.
- Stenstrom, A. (1994). *An Introduction to Spoken Interaction*. Longman, London and New York.
- Swerts, M. (1998). Filled pauses as markers of discourse structure. *Journal of Pragmatics*, 30:485–496.
- Tannen, D. (1985). Silence: Anything but. In Tannen, D. and Saville-Troike, M., editors, *Perspectives on Silence*, pages 93–111. Ablex, Norwood, NJ.
- Viney, P. and Viney, K. (1996). *Handshake*. Oxford University Press, Oxford, UK.
- Voss, B. (1979). Hesitation phenomena as sources of perceptual errors for non-native speakers. *Language and Speech*, 22:129–144.