

# Disfluent Speech

(2,418 words)

## Article Table of Contents

1. [1. Introduction](#)
2. [2. Structural and Prosodic Patterns of Disfluent Speech](#)
3. [3. Filled Pauses and Speech Repairs in Mandarin Conversation](#)
4. [4. Concluding Remarks](#)
5. [Bibliography](#)

## 1. INTRODUCTION

The study of disfluent speech is often regarded as peripheral in linguistic research. Nevertheless, a considerable proportion of conversational speech is disfluent. If we want to understand how human speech operates, it is indispensable to study how and when disfluent speech is produced. Explorations into linguistic representations of disfluent speech have attracted special attention not only from linguists and psychologists, but also from computer scientists, as interest in spoken dialogue systems has dramatically increased due to practical applications concerning human speech processing and understanding (Lin *et al.* 2009). Knowledge about the structural patterns and prosodic cues of disfluent speech produced in an interactional context is required to understand the way disfluent speech is produced by humans and the way it should be processed by automatic systems.

In providing the background for speech monitoring, Levelt (1983) distinguished three main types of speech repairs: (1) *Do I want to say this now?*; (2) *Do I want to say it this way?*; (3) *Am I making an error?*, respectively calling them D- (different), A- (appropriate), and E- (erroneous) repairs, in which speakers say something to make their earlier delivered message different, appropriate, or correct. In terms of speech as a sequence of events, Shriberg (2001) proposed six disfluency types: repetition, deletion, substitution, insertion, articulation error, and filled pauses. Hieke (1981) proposed a system of hesitations: *stalls* (silent pauses, filled pauses, prospective repeats, syllable prolongation) and *repairs* (false starts, retrospective repeats or bridging).

Different focuses lead to different taxonomies for disfluent speech. On the whole, four main types of disfluent speech can be distinguished, though not mutually exclusive: pauses, hesitation, repairs, and false starts. As shown in example (1) (Lin *et al.* 2009:194), *jìnkǒu* 進口 'import' is the to-be-repaired item, *chūkǒu* 出口 'export' the repaired item, and *uhn* (transcribed here as 嗯) is a filled pause that usually has a function of hesitation (COP = copula, QUES = question particle). Levelt (1983) named everything that occurs between the to-be-repaired and the repaired items “the editing term”. The editing term can be a filled pause such as *uhn* in the example, a silent pause or lexical items such as *you know* and *well* in English.

1. 是進口嗯出口嗎？

Shì jìnkǒu uhn chūkǒu ma?

COP import filler export QUES

‘Do (you) import uhn export (products)?’

Silent pauses in disfluent speech differ from rhetorical pauses that are intentionally made to form a sensible prosodic segmentation of discourse. A silent pause indicates interruption, whereas filled pauses indicate hesitation. Both silent pauses and hesitation can occur in the middle of grammatically correct spoken utterances, or between the to-be-repaired and the repaired items. Hesitation with the function of stalling is often realized as a prolonged form of lexical words, or as filled pauses (also called fillers) produced by vocalizing (1) a schwa-like vowel with an optional nasal consonant, e.g., *uh* and *uhn*, (2) lexicalized items such as *I mean* in English (Shriberg 2001), or (3) demonstrative filled pauses such as *nà* 那 'that' in Chinese (Zhao and Jurafsky 2005).

Repairs and false starts share similar structures, but differ at the linguistic level (Chui 1996, Couper-Kuhlen 1992, Levelt 1983, Schegloff *et al.* 1977, Shriberg 1994, Tseng 2006). Speech repairs are usually regarded as an utterance-internal phenomenon, as the correction of an error is often initiated at a phrasal or word boundary (Levelt 1983, Tseng 2006). False starts refer to the sentence level, where the speaker abandons the unfinished utterance and starts with a new one. The repaired item repeating part or all of the to-be-repaired item is called recycling, as shown in example (2), where the morpheme *yī* 一 'a' is repeated (Tseng 2006:95).

2. 可能有一陣一段時間。

Kěnéng yǒu yízhèn yíduàn shíjiān.

possibly there is a period a piece time  
'Possibly there is a period a certain time.'

When the entire to-be-repaired item is repeated, it is also called a repetition, as illustrated in example (3) (Tseng 2006:95).

3. 想說明天明天去。

Xiǎng shuō míngtiān míngtiān qù.

want say tomorrow tomorrow go

'I'm planning to go tomorrow tomorrow.'

Concerning Mandarin speech repairs, studies have been done accounting for different discourse perspectives. On the basis of 241 instances of self-repair produced in Mainland Mandarin conversation Tao (1995) proposed three linguistic mechanisms: recycling, replacement, and restructuring. Chui (1996) suggested six patterns of repair based on 463 instances of self-repair in Taiwan Mandarin conversation: repetition, completion, replacement, addition, reordering, and abandonment. Zhang (1998) investigated how self-repairs are organized within speaker turn managements (i.e., how conversational partners negotiate who talks when) in a dataset of 13 hours of calls to radio programs in Mainland Mandarin.

## 2. STRUCTURAL AND PROSODIC PATTERNS OF DISFLUENT SPEECH

Conversational analysts are concerned with the role disfluent speech plays in human speech communication, in particular in the context of discourse organization (Schegloff *et al.* 1977). Utilizing speech repair data collected from production experiments, structural patterns and prosodic realizations were systematically examined by Levelt and Cutler (1983) for Dutch. They found that prosodic marking with a noticeable increase or decrease in fundamental frequency (F0), duration, and amplitude is used to distinguish the meaning of the actual repair and the to-be-repaired words, when necessary. A similar prosodic pattern was also found by Couper-Kuhlen (1992) for English, where repair sequences were characterized by an increase in tempo. Such observations imply that phonology can be interactional, facilitating a kind of prosodic contextualization and thus also reinforcing the processing of repair. Table 1 summarizes the structural and prosodic cues of each disfluency type that have been mentioned in the non-Chinese literature (Clark and Wasow 1998, Levelt and Cutler 1983, O'Shaughnessy 1993, Shriberg 1994). These patterns mostly also apply to Chinese (Chen 2011, Tseng 2006, Zhao and Jurafsky 2005), except that

little has been researched regarding repairs of lexical tones that may affect the F0 patterns.

Table 1. Structural and prosodic cues for various disfluency types

Types	Pauses	Hesitation (Prolongation/filled pauses/discourse markers)	Repairs	False starts
<b>Structural pattern</b>	None	None/filled pauses/discourse markers	Repeated words/part of speech	Abandoned utterances
<b>Prosodic clues</b>	Silence	Lengthening/low F0/flat F0 contour/duration	Pitch reset/duration/amplitude change	Pitch reset/duration

The patterns above may occur in simple or combined forms. Chen (2011) proposed six prosodic patterns of Chinese recycling repairs in terms of duration, amplitude and pitch height, each pattern with a specified function in interactional context. As shown in example (4), adapted from Chen (2011:70), the function of a longer to-be-repeated word with a cut-off followed by a repeated word at the same pitch height indicates continuation of the current discourse topic (PN = personal name, FP = filled pause, NEG = negation, AUX = auxiliary, QUES = question particle, % = cut-off).

4. Y: 小 C 不要睡覺喔。

Xiǎo C bù yào shuìjiào o.

little P N not AUX fall asleep FP

C: 不會(sound)我我回去洗澡可以嗎?

Bú huì (sound) wǒ% wǒ huíqù xǐzǎo kěyǐ ma?

not AUX 1SG 1SG return take a shower AUX QUES

Y: 不行。

Bù xíng.

not work

Y: '(To C) Little C, don't fall asleep!'

C: '(I) won't.' '([sound])(can) I% I go back taking a shower, is it all right?'

Y: 'No.'

However, no matter how complex a repair is, the repeated items are in general shorter than the to-be-repeated items. Clark and Wasow (1998)

suggested that speakers repeat words to restore continuity to the delivery of the constituent they were initiating. Since the part carrying known information is processed as quickly as possible in an interactive discourse, the repeated words in a repetition or repair are generally produced faster (O'Shaughnessy 1993; Shriberg 1994).

Computational linguists have attempted to automatically detect interruption points in Chinese spontaneous speech by using acoustic-prosodic features of silent pauses and speech repairs as well by shallow parsing that identifies repeated words and word categories. But so far only marginal improvement has been accomplished in the performance of automatic speech recognition systems (Lee and Chen 1997; Lin *et al.* 2009).

Concerning lexical tones in Mandarin, Tao *et al.* (1999) studied 120 instances of tone-choice repairs of Mainland Mandarin. Tone errors are mainly caused by a wrong tone selection. For example, the tone of the first syllable in *qíshí* 其實 'in fact' was wrongly pronounced as a high level tone [55] instead of the rising tone [34] in *qi55 qi34qi34*, as illustrated by Tao *et al.* (1999:272). Or a wrong tone sandhi form is taken, as in the example *zhao211.5 zhao35 ji213'e tong35xue34* for *zhǎo jǐge tóngxué* 找幾個同學 'find a few students' (Tao *et al.* 1999:269). The first *zhao* was realized with a low tone (211.5, as noted by Tao *et al.* 1999) that is the realization form of the third tone when followed by a non-third tone (unlike here, since *jǐ* has third tone). Erroneously lengthened vowels also led to tone errors such as *nii suàn nǐ suàn* for *nǐ suàn* 你算 'you count' (Tao *et al.* 1999:277), in which the lengthened *ni* led to a prolonged rising tone 2244. Tao *et al.* (1999) also found that tone errors are mostly immediately repaired in the subsequent word.

### 3. FILLED PAUSES AND SPEECH REPAIRS IN MANDARIN CONVERSATION

In the Taiwan Mandarin Conversational Corpus (TMC, Tseng 2013), 6,338 simple filled pauses (*uhn, uh, mhm*) and 16,516 demonstrative filled pauses *nà* 那 or *nàge* 那個 were produced, at ratios of 15.6 and 40.7 times per 1000 words for simple and demonstrative filled pauses, respectively. The demonstrative filled pauses *nà* and *nàge* literally mean 'that' and 'that one', but often in colloquial speech, the original semantics is lost, while the pragmatic function becomes more salient (Tseng 2006, Zhao and Jurafsky

2005). Zhao and Jurafsky (2005) reported the results of filled pauses produced in Mainland Mandarin Chinese telephone conversation. The most frequent filled pauses *uh* and *mhm* (notated as *mm* in Zhao and Jurafsky 2005) occurred 2.55 and 1.46 times per 1000 words, respectively, while *nàge* and *zhège* 這個 occurred 4.51 and 2.17 times per 1000 words, respectively. The numbers of filled pauses were thus much smaller in the Mainland Mandarin corpus than in the Taiwan Mandarin corpus, presumably due to differences in Mandarin dialect and corpus setting. The Taiwan Mandarin corpus consists of face-to-face conversations, whereas the Mainland Mandarin corpus contains telephone conversations.

In Tseng's (2006) data covert speech repairs (repetitions) are produced more frequently than repairs that actually "correct" a prior utterance. Simple repetitions mostly involve only one phrase, while the most frequently repaired syntactic categories are nouns (42%), adverbial phrases (26%) and verbs (23%), whose lexical coverage in spoken Táiwan Mandarin are 22%, 20%, and 23%, respectively (Tseng 2006; Tseng 2013). This comparison suggests that nouns are more likely to be repaired than the other categories in Mandarin. Tseng (2006) also found that repeated and corrected words in the repair part are marked by a faster tempo, a shorter duration, and a weaker intensity. The pitch values of the onsets of the to-be-repaired and the repair parts are not significantly different in Tseng (2006), indirectly suggesting a pitch reset (as a null hypothesis is difficult to be proved statistically).

#### 4. CONCLUDING REMARKS

We use the term "disfluent speech", but being disfluent is nothing but "human" (Shriberg 2001). It is common to observe these so-called ungrammatical sequences or performance in spontaneous, unprepared conversation, which rarely result in understanding problems for humans. But in automatic speech systems, it is extremely difficult to generate a cleaned up version from disfluent speech data. Maybe this is because we do not yet know enough about disfluent speech to make an automatic speech system "human-like" enough. More studies on Chinese tone errors and repairs in conversational speech are needed to shed light on how tones are associated with the segmental structure of Chinese spoken words. The prosodic patterns concerning F0 contours and duration in disfluent speech in Chinese may vary, if tone repairs are taken into account.

## Bibliography

Chen, Helen Kai-yun, “Sound Patterns in Mandarin Recycling Repairs”, dissertation, University of Colorado, 2011.

Chui, Kawai, “Organization of Repair in Chinese Conversation”, *Text* 16, 1996, 343–372.

Clark, Herbert H. and Thomas Wasow, “Repeating Words in Spontaneous Speech”, *Cognitive Psychology* 37, 1998, 201–242.

Couper-Kuhlen, Elizabeth, “Contextualizing Discourse: The Prosody of Interactive Repair”, in: Peter Auer and Aldo di Luzio, eds., *The Contextualization of Language*, Amsterdam: John Benjamins Publishing Company, 1992, 337–364.

Hieke, Adolf, “A Content-Processing View of Hesitation Phenomena”, *Language and Speech* 24, 1981, 147–160.

Lee, Yue-Shi and Hsin-Chi Chen, “Using Acoustic and Prosodic Cues to Correct Chinese Speech Repairs”, *Eurospeech* 97, 1997, 2211–2214.

Levelt, Willem, “Monitoring and Self-Repair in Speech”, *Cognition* 14, 1983, 41–104.

Levelt, Willem and Anne Cutler, “Prosodic Marking in Speech Repair”, *Journal of Semantics* 2, 1983, 205–217.

Lin, Che-Kuang, Shu-Chuan Tseng and Lin-Shan Lee, “Spontaneous Mandarin Speech Recognition with Disfluencies Detected by Latent Prosodic Modeling (LPM)”, in: Shu-Chuan Tseng, ed., *Linguistic Patterns in Spontaneous Speech*, Monograph Series A25, Institute of Linguistics, Academia Sinica, 2009, 193–212.

O’Shaughnessy, Douglas, “Analysis and Automatic Recognition of False Starts in Spontaneous Speech”, *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Vol. II, 1993, 724–727.

Schegloff, Emanuel, Gail Jefferson and Harvey Sacks, “The Preference for Self-Correction in the Organization of Repair in Conversation”, *Language* 52, 1977, 361–382.

Shriberg, Elizabeth, “Preliminaries to a Theory of Speech Disfluency”, dissertation, University of California, 1994.

Shriberg, Elizabeth, “To ‘errrr’ Is Human: Ecology and Acoustics of Speech Disfluencies”, *Journal of the International Phonetic Association* 31, 2001, 153–169.

Tao, Liang, “Repair in Natural Conversation of Beijing Mandarin”, *The YuenRen Society Treasury of Chinese Dialect Data* 1, 1995, 55–77.

Tao, Liang, Barbara Fox and J. Gomez de Garcia, “Tone-Choice Repair in Conversational Mandarin Chinese”, in: Barbara Fox, Dan Jurafsky and Laura Michaelis, eds., *Cognition and Function in Language*, Stanford: Stanford Center for the Study of Language and Information, 1999, 268–281.

Tseng, Shu-Chuan, “Repairs in Mandarin Conversation”, *Journal of Chinese Linguistics* 34, 2006, 80–120.

Tseng, Shu-Chuan, “Lexical Coverage in Taiwan Mandarin Conversation”, *International Journal of Computational Linguistics and Chinese Language Processing* 18, 2013, 1–18.

Zhang, Wei, “Repair in Chinese Conversation”, dissertation, University of Hong Kong, 1998.

Zhao, Yuan and Dan Jurafsky, “A Preliminary Study of Mandarin Filled Pauses”, *Disfluency in Spontaneous Speech (DiSS'05)*, 2005, 179–182.

[Cite this page](#)

Shu-Chuan TSENG. "Disfluent Speech." *Encyclopedia of Chinese Language and Linguistics*. General Editor Rint Sybesma. Brill Online, 2016. [Reference](#). BRILL demo user. 25 April 2016 <[http://referenceworks.brillonline.com/entries/encyclopedia-of-chinese-language-and-linguistics/disfluent-speech-COM\\_000095](http://referenceworks.brillonline.com/entries/encyclopedia-of-chinese-language-and-linguistics/disfluent-speech-COM_000095)>

Last online update: March 2016

First appeared online: 2015