

# Praat-assisted Comparative Study on Disfluencies in C-E Interpreting

Yang Xiaofeng<sup>1</sup>, Zhang Menglei<sup>1</sup>, Zhang Zhaoqing<sup>1</sup>, Zeng Ashan<sup>2,\*</sup>

<sup>1</sup>Department of Foreign Languages, Northwest A&F University, Yangling, China

<sup>2</sup>School of Translation, Xi'an International Studies University, Xi'an, China

## Email address:

Yangxiaofeng78@nwsuaf.edu.cn (Yang Xiaofeng), 176969205@qq.com (Zhang Menglei)

\*Corresponding author

## To cite this article:

Yang Xiaofeng, Zhang Menglei, Zhang Zhaoqing, Zeng Ashan. Praat-assisted Comparative Study on Disfluencies in C-E Interpreting. *International Journal of Applied Linguistics and Translation*. Vol. 7, No. 2, 2021, pp. 76-82. doi: 10.11648/j.ijalt.20210702.17

**Received:** January 14, 2021; **Accepted:** March 1, 2021; **Published:** June 25, 2021

---

**Abstract:** Intervening the process of interpreting, tone group segmentations, as the phonological embodiment of the interpreters' information perception and processing, produce dis-fluent paralinguistic utterances like pauses, lengthening, and pitch resetting etc, which impact on the fluency of the information and expression. In order to study how these non-verbal behavior impose on the disfluencies and explore new criteria to assess interpreting quality technically, thus training students' interpreting skills, the study applied Praat to study on the students' disfluency phenomenon in C-E interpreting from lexical, phrasal and sentential levels and tones and tone groups segmentation in student's C-E interpreting disfluent utterances in PACCEL 2012 with the interpreting of the 2010 SC Press conference as CK. The results showed that, compared with the professional interpreters, student interpreters' utterance were endowed with the external and inner-word breaks, broken phrase segmentation, and excessive tone group segmentations in sentences.

**Keywords:** Interpreting Disfluency, Tone Groups Division, PACCEL, Praat

---

## 1. Interpreting Disfluency

Interpreting is a process of converting information from one language to another to realize information equivalence and assist instant understanding between relevant communicating parties (Hu, 1988). Interpreting is never confined to mechanical transformation of original language code, but more on information spreading and exchanging communicative activities in particular situations (Li, 2004; Liu, 2005). To obtain efficient communication, interpreters should ensure their performance effectively help for the audience's understanding and perception, and this urges a lot on study the quality of interpreting, especially the standard for us to judge qualified interpreting.

Faithfulness, expressiveness and elegance are widely accepted in translating quality assessment. However, interpreting is different from translating. It's a time-limited communicative activity. Due to the information input, bilingual syntactic mismatch, heavy load of tasks at same time and great pressure in coordination, utterance like pause,

broken, lengthening, repeating, hesitating or reversing might frequently occur in the process of interpreting. Li (1987) put forward the standard of "Fluency, accuracy, and quickness" into the assessing criterion of qualified interpreting and from then on in the interpreting quality evaluation system, fluency is regarded as one of the key elements in assessing the interpreting quality as well as interpreters' professionalism.

Studies showed that not all the hesitating or pausing disfluencies in interpreting are mere errors or results of inadequacies. In the interpreting process, interpreters will purposely segment the tone group to affect listeners' transmission, understanding and perception of the information.

The study on disfluency, specifically hesitation and pause phenomenon, has propelled in the middle of 20th century (MacLay & Osgood, 1959). Fromkin (1998) and Levelt (1983) both contributed a lot as the former researched speech errors from the perspective of grammar and the other from the purpose of self-repairing. Fox Tree (1993) advanced his comprehension of speech disfluencies. Lickley (1994) is devoted to addressing the problem of detecting disfluency in

spontaneous speech through figuring out those detecting points and the accompanying acoustic and prosodic cues and five perceptual experiments were undertaken to obtain the data in his paper. As for the detection issue, Gabrea & O'Shaughnessy (2000) had a more specific research concerning the detection of filled pauses in spontaneous conversations. Shriberg (1994) provided evidence that disfluencies have a regular trend in many dimensions.

In Chinese-English interpreting, Yang Jun [9] studied the disfluency phenomenon of oral output and stressed on three typical disfluency phenomenon--- pauses, repetition and self-repair. He argued that filled pauses are often included in many disfluencies and sometimes even considered as a sign of self-repair. These disfluent utterances appear more at the beginning of a sentence and the length increased when sentences are longer. In principle pauses before low-frequency utterance outnumbered those before high-frequency, similar to the finding by Klatt (1980).

Whilst actual interpreting, the clients expect to receive complete and accurate information from the interpreters. Therefore, it's better to go without pauses in interpreting (Herbert, 1982), and the idea of "disfluency" turned to be a main standard judging interpreting quality (Buheler, 1986). The interrelationship between disfluency and interpretation quality were surveyed among clients (Macia, 2003) and the result showed that information and utterances fluency ranked the 3<sup>rd</sup> and the interpreting consumers held great expectations on smoothness and fluency of the utterance by interpreters.

Due to its nonverbal paralinguistic features, study on disfluencies needs to be assisted by certain visible and analytical schemes, and in the present study, we set tone group as the access to decode disfluencies from the perspective of Phonology and pragmatics. Tones, further defined to include tonality, tonicity, and tone, is not only the change of pitches in production, it is also a combination of tones, syntax, and information (Halliday, 1975). Tone group division refers to the pitch pattern of oral English in terms of suprasegmental Phonology (O'Connor & Arnold 1973:1). To specify the pattern, Cruttenden (1997) put forward four external features concerning tone group division---pause (the break between two words with the minimum duration of 0.3s (Raupach, 1980)), anacrusis (unstressed syllables from the beginning of a sentence), lengthening (lengthened stressed syllable, or a substitute for pause), and pitch resetting. Study on tone group division in China started from research on reading patterns, such as the intonation pattern and acoustics of Chinese English learners [11], English tone grouping pattern in reading (Zhu, 2014).

## 2. Research Process

Mega corpora have become a reality with the advancement of technology and computer data processing in the 1990s. From then on, miscellaneous corpora in various languages come into being. More than 10 years ago, an article published

by Schlesinger has proposed corpus-based interpreting study (CIT) as a branch of corpus-based translation studies (CTS). Earliest studies on CIT can date back to the late 20th century (Shlesinger, 1998), and now more professional CIT corpora has been set up for academic and training purposes (Li & Li, 2010). On top of the lists is the Simultaneous Interpretation Database of Nagoya University in Japan which consists of the recording and transcription of 182 hours' interpreting, the largest of its kind in the world. Besides, University of Bologna established the European Parliament Interpreting Corpus (EPIC) (Monti et al., 2005), a multilingual parallel corpus of English, Italian and Spanish. In China, CIT study started lately, marked by Parrellel Corpus of Chinese EFL Learners (PACCEL), Chinese English Conference Interpreting Corpus (CECIC) and so on [10].

In order to study how the non-verbal segmentation of tone groups would impose on the disfluency in interpreting utterances, and detect how the professional interpreters manage their tones, and why their rhymes are much more euphonic, this paper applied Praat (Paul Boersma & David Weenink) and compared the tone group division between learners and professional interpreters in the two corpora of PACCEL and professional interpreting on Chinese Prime Minister Press Conference 2010.

At the convenience of comparison, a sentence by 15 boys and 15 girls from C-E interpreting test in PACCEL 2007 has been selected to conduct contrastive analysis against another similarly-structured sentence by professional interpreter.

Import the object sound to Praat software and create a corresponding Text Grid with 4 tiers, namely Sentence, Words, V&C, and B (see Figure 1). The "Sentence" tier includes the whole sentence, while the "Words" tier segments words according to the pitch of every word. "V&C" tier labels phonogram of each word sound and the "B" tier was a point tier to mark tone groups. Generally, number "4" or sometimes "5" is used to be the mark of division of tone groups.

To assist the comparison, two sound pitch wave lines were drawn on the lexical level, of which the blue one stands for professional interpreters' pronunciation, while the black one stands for that of the students'. Number of "5" and "4" were used to mark the segmentation on the phrasal and tone group divisions on sentential levels.

In the experiment, a sentence from C-E interpreting Test of 2003 English Major Band-8 has been chosen. The original Chinese sentence is "Women (we) Juban (hold) zhege (this) Zhanlanhui (exhibition) DE ('s) Mudi (purpose) shi (BE) Jiaqiang (enhance) zai (PREP) zheyi (this) Lingyu (domain) DE ('s) Guoji (international) Jiaoliu (communication) yu (and) Hezuo (cooperation)."

To ensure reliability and comparability, overlapping rate of students' interpreting with that of the professional is calculated and the results showed that at least 90% of words by students are included in the interpreting by the professional interpreter (Table 1).

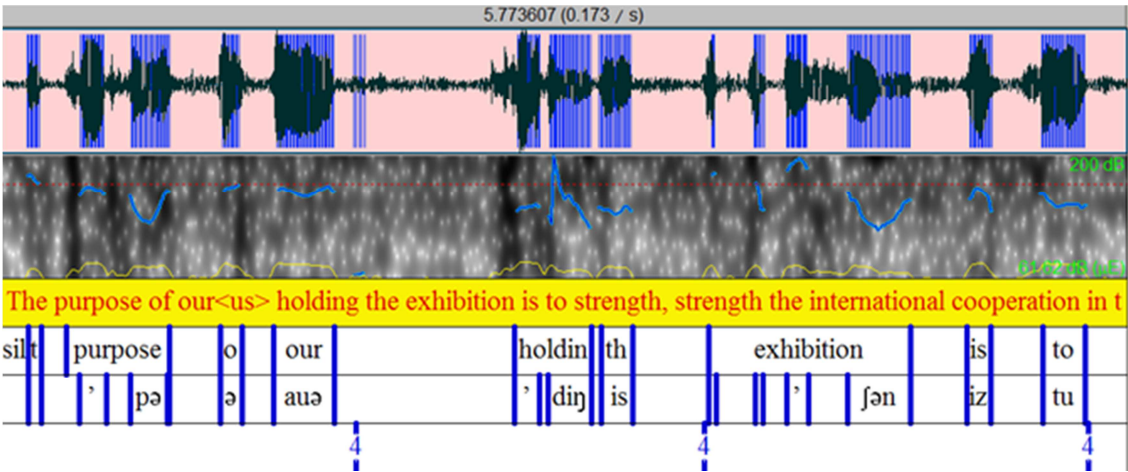


Figure 1. Text Grid Tiers.

Table 1. Overlapping Rate of words by students.

Student Number	Overlapping Rate	Student Number	Overlapping Rate	Student Number	Overlapping Rate
1	94.1%	11	100%	21	97%
2	91.6%	12	90%	22	96%
3	92.2%	13	90.9%	23	94.1%
4	90.0%	14	90%	24	100%
5	90.4%	15	94.7%	25	95.4%
6	94.7%	16	92.5%	26	93.5%
7	95.2%	17	90.9%	27	94.1%
8	100%	18	95.8%	28	95%
9	93.7%	19	94.4%	29	95.4%
10	94.4%	20	90%	30	92%

3. Research findings

Pitch lines at lexical level  
To compare the pitch lines, we selected the word of

“purpose” in the interpreting and established its pitch lines. Although there revealed slight differences, students’ lexical pitch lines can be classified into two kinds, as is in Figure 6a and 6b.

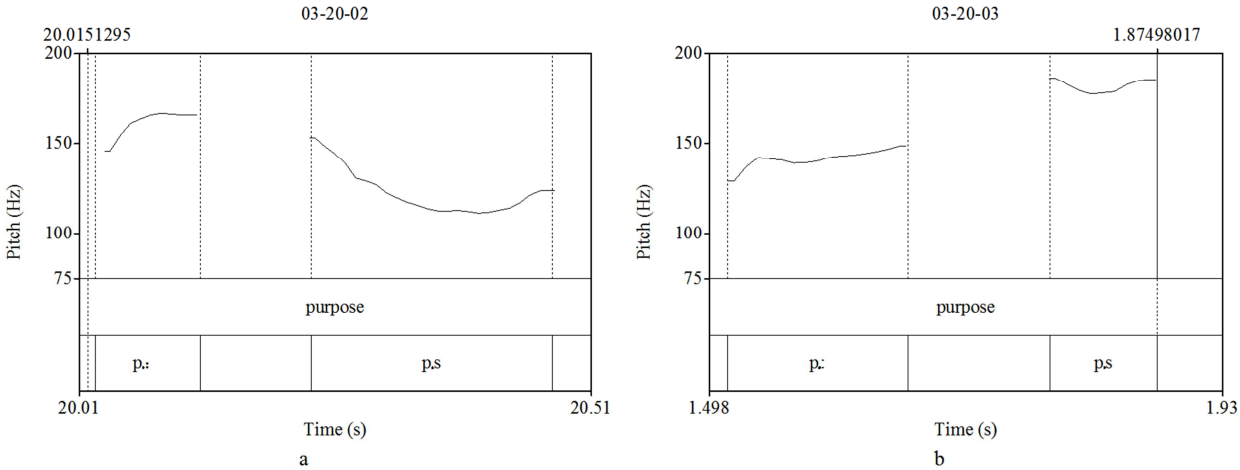


Figure 2. Pitch Lines of “PURPOSE”.

As can be seen from the two graphs, students have some breaks when pronouncing the word “purpose”. At the second syllable of the word, some students produced a higher syllable pitch line compared with the first syllable (2a); and some, produced a lower one (2b). When 2a trended to decline in pitch, 2b rose up.

For reference, a comparative graph with word utterance by

professional interpreter, Ms. Zhang Lu, is compiled to elaborate the differences.

A noteworthy phenomenon is that many words by student interpreters have breaks in their sound (Figure 3). The paper collected words without breaks by Zhang Lu, and compared the sound wave of the words.

Judging from the sound waves of words “purpose, holding,

field” by Zhang Lu (Figure 4), we can easily see that professional interpreters’ utterances run without any breaks. All sounds are consecutively connected and the sound waves had ups and downs, a good indication of rhymes. They went like natural sound flows from one syllable to another. Yet the waves of students’ sounds were quite broken. As is collected in figure 3, more than 70% of among the tested 30 students had breaks in those three words.

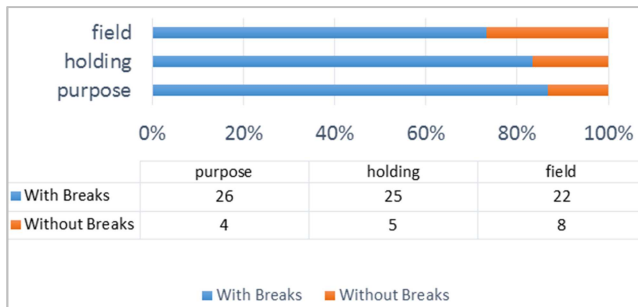


Figure 3. Words With Breaks from Students.

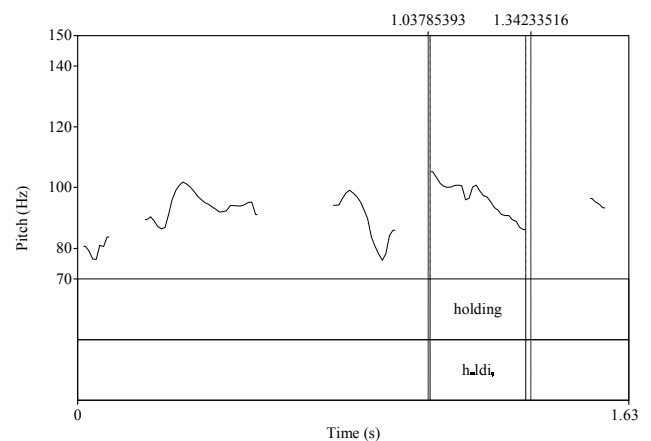
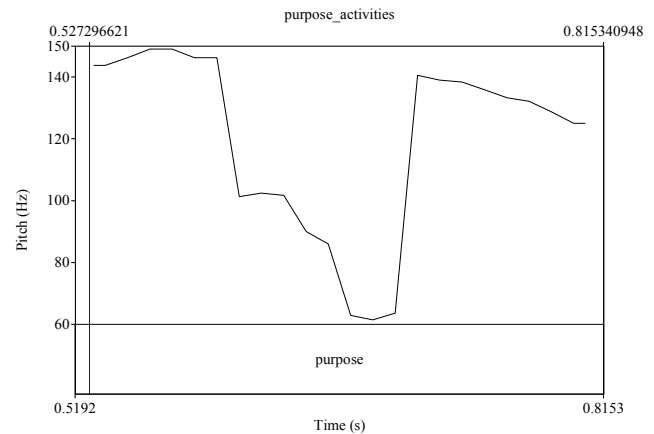
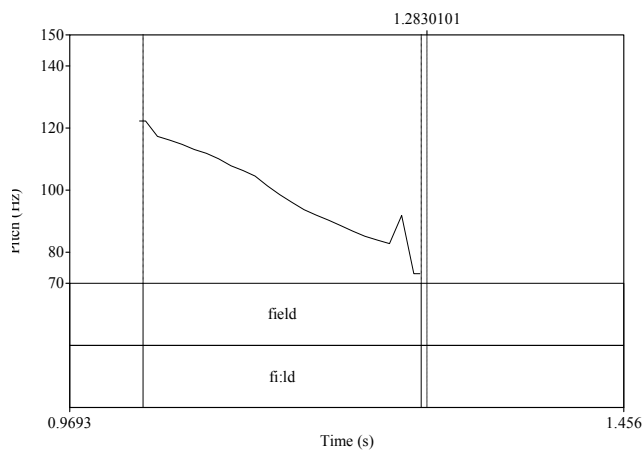


Figure 4. Sound waves of words by professional interpreter.

When there’s a stress in the second syllable or hinder part of a word, or when there’s a plosive sound, the air flow in sound production would not be smooth and sustaining. If there’s no stress or if the stress is in the first syllable of a word, the sound wave of the word should be without any breaks.

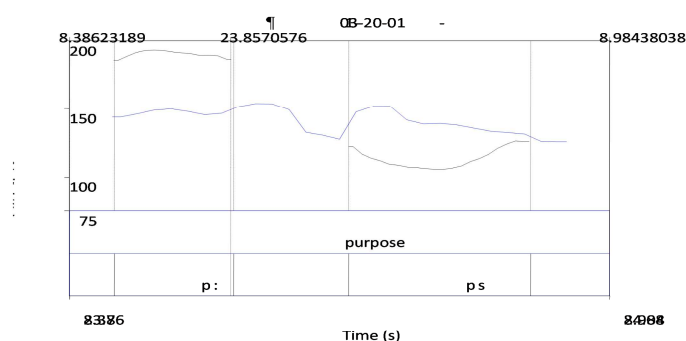


Figure 5. Student-professional interpreter pitch lines of “PURPOSE”.

Figure 5 contrasted the pitch lines of the word “purpose” by Zhang Lu colored in blue, and student 1 colored in black. The phonogram of the word PURPOSE is as /'pɜː pəs/. There’s no stress in the second syllable so there should be no breaks in the sound wave, just as the blue line flows. However, in the

students’ utterance, the breaks indicated that students have segmented the sound, and thereby resulted disfluent pronouncing at the lexical level.

Pitch lines of the word VICTORY might display similar breaks by the professional interpreter. However, a closer

observation on the graph will find out that actually the pitch line of VICTORY is not consecutive. The phonogram of the word “victory” is /vɪkt (ə) rɪ/. There’s no stress in the second or third syllable of the word, thus the second sound /t ə/ should be on the same level with the sound /vik/, however it does not. It proves the seemingly breaks in the word VICTORY is pitch resetting inside a word. Pitch resetting which often occurs at the boundaries between prosodic units can also occur inside the word structure. In contrast, breaks inside the students’ production of sound stopped the flow of pitch line and can not be regarded as pitch resetting.

#### Information units at phrasal level

Information unit is the division of target information and is always taken as the working unit in translation and interpreting. When interpreting, interpreters usually divide the source language information into segments, among which phrases are commonly adopted. For example, in “I’ll call you back as soon as I got home”, phrases like “as soon as”, “call you back” and “got home” will be normally clustered as a whole information unit. Information units not only worked in semantic decoding and encoding, they could be read or produced as a whole in spoken process. The general tendency is that the bigger an information unit is clustered, the more fluent the speech might sound. To minimize the errors of information unit detection, the study simply take phrasal

division as the index of measuring.

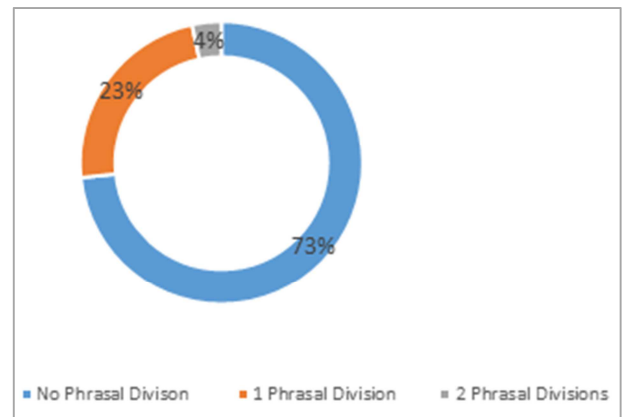


Figure 6. Students' Phrasal Divisions.

According to Figure 6, more than 70% of the 30 students did not divide any phrasal units in their sentence and they linearly interpreted the sentence word by word. 23% had 1 phrasal division and 4% 2 phrasal divisions in their interpreting, which indicated that student interpreters seldom adopted phrasal division, but interpreted word by word, which can cause intervals in speech flow.

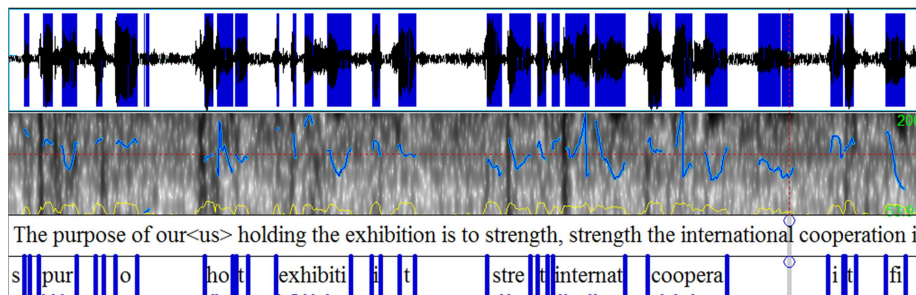


Figure 7. Phrasal divisions by student interpreter.

Figure 7 is a case of phrasal division by student No. 1. It can be seen the 17 words of this sentence seems independent from each other in terms of the pitch line and “V&C” Tier. There’s no phrasal division in the sentence, but breaks existed even inside a single word.

For comparison, the study cited an equivalent case by professional interpreter (2010 PM Press Conference). The original Chinese sentence is interpreted as “Not only China

*but also many countries in the world have planned to hold diverse forms of commemoration activities. The purpose of these activities is to firmly bear in mind that the lessons gained from the past and ensure that that kind of history will never repeat itself. The purpose is to uphold the outcomes of the World War II and the post-war international order and international laws to maintain enduring peace of mankind.”*

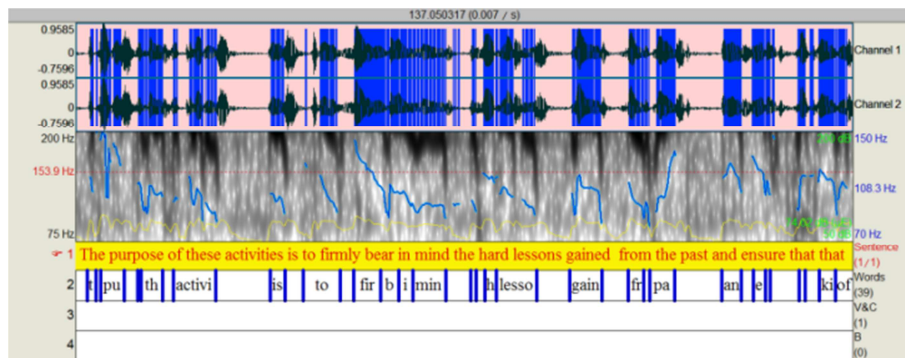


Figure 8. Phrasal divisions by professional interpreter.



At the sentential level, Zhang has two phrasal divisions in the first sentence consisting of 19 words, 4 divisions in the second sentence with 29 words and 3 in the third with 25 words. The following graph shows the phrasal division in the second sentence by the professional interpreter.

In the graph, the pitch line shows that there's no intervals among the words "firmly bear in mind", "from the past" and "that kind of". All words had been attached with each other to form a whole unit of sound and information.

In summary, different from the interpreting strategy and its caused phrasal division by professional interpreter, student interpreters mostly interpreted word by word, and they

seldom or did not combine or divide sounds unit, also regarded as information unit, in the process of interpreting, which therefore led to the broken flow of speech.

**Tone groups division at sentential level**

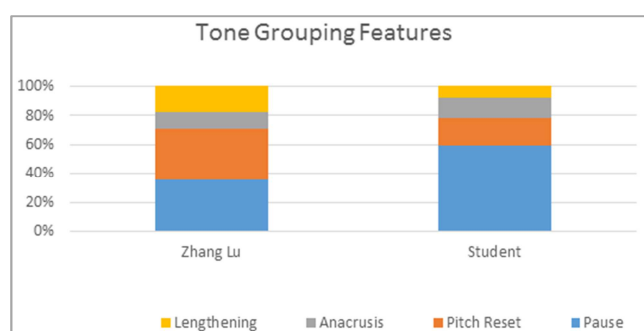
Tone group division refers to the segmentation of a long sentence according to the flow of speech. Tone group also stands for the information unit (Halliday, 1994/2000); it's the perception and information organization of interpreters (Chen, 2006). Thus, the division of tone groups displays the interpreters linguistic perception of the source language information, furthermore affects the receivers' perception to the interpreted target language, usually, in chunks of information.

**Table 2.** Tone Group Numbers.

	Mean Number of Words	Sentence Number	Mean Tone Groups
Student Interpreter	20	30	12.3
Professional Interpreter	24	3	3

Table 2 showed that although having the similar words number, the student interpreters averagely divided three times more tone groups than the professional interpreter. This can be partially explained by the higher linguistic competency of professional interpreters, also could be the reason to explain the cohesive and consecutive speech flow and sounds waves of the better interpreters. Compared with Zhang, the professional interpreter, student interpreters spend longer time for sentence of similar length and had more pauses in process of interpreting.

As mentioned above, Tone groups had four distinctive external features: pause, anacrusis, syllable lengthening, and pitch reset of unaccented syllable. Although these four features may be either purposely produced or passively uttered, pauses break the fluency of sounds and meaning, while the other three may as well contribute to a pleasant pronouncing effect if properly managed. To compare the distribution of external features applied by different level of interpreters, the study classified the ton groups of professional interpreter and that of the student interpreters (Figure 9).



**Figure 9.** Student-Professional Interpreter Tone Groups Features.

As is seen in Figure 8, pauses and pitch resetting are the top two external features adopted by both the student and professional interpreters, about 70% for professional and over 80% for students of the total. However, compared with Zhang Lu who used 30% of pauses and 40% of pitch resetting in her total tone groups division, the student interpreter mostly relied on pauses, nearly 60%, indicating that pauses by students might

help to explain the bigger number of tone groups and broken and stumbling sound flow of the interpreting. More pauses made the speech flow disconnected and thereby affect the receivers' perception of target information. Noteworthy is that pitch resetting relates closely to the tones of the sentence, and can help realize a rhythmic sound flow in interpreting, which may to some extent contribute to the favoring sound and fluency of professional interpreters.

There seemed no big difference for using of anacrusis, which might be due to the limited number of samples. While for lengthening, professional interpreters used two times of this more than the student interpreters.

## 4. Conclusion

Based on the above indexes and their contrastive analysis, it can be safely concluded that both professional interpreters and student interpreters would have pauses in their speech. Thanks to fewer pauses, professional interpreters produced fluent and smoother flow of speech.

At lexical level, student interpreters have breaks between two words or phrases, even inside a word. Although it is not applicable to distinguish differences of the word with or without breaks, the utterances still affects the auditory effectiveness of their interpreting.

At the phrasal level, students seldom have phrasal divisions and most of their interpreting happened word by word, which not only portrayed their inadequacy in linguistic chunks for processing interpreting tasks, also explained the fact why they divided more tone groups and therefore effected disfluent speech flow.

At the sentential level, tone groups division revealed the interpreters' perception and decoding of the source language information. The more tone groups the interpreter divide, the smaller units he separated the original information, and the more scattered the perception of information chunks of receivers might be. In contrast, few tone groups might stand for bigger information unit, more cohesive ideas and more fluent information flow. External tone group features analysis

discovered that pauses could be the reason for disfluencies in interpreting and other features, or strategies could help to maintain the unity and cohesiveness of idea, and contributed to a fluent and favorable auditory effectiveness in interpreting.

The study concludes that disfluencies in process of interpreting, though seem to be articulation or sound production errors, had its root causes in linguistic perception and application. Based on the analysis, it can be concluded that quality of interpreting correlates positively and significantly with tone group divisions. Therefore, in Chinese to English interpreting, trainers of interpreters can start from tone groups, or chunking information practicing, help them attain better command of proper tone group division and phrase divisions, and cultivate fluent interpreters from bilingual information chunking, phrasal building and tone group division.

The whole research could be more elaborative if the sample could be bigger, especially those of professional interpreter, if the student interpreter selection could be more representing, and the labeling and analyzing software could be more purposely applied in further study on the phenomenon of pauses. We do feel short of findings and limitation on research methods due to technical ignorance. PACCEL and Praat are very good data base and instruments, but the proposed empirical study on the acoustic and articulation performance of interpreters will surely promise more meaningful and targeted guidance to millions of interpreting learners in China.

## Acknowledgements

This research is supported by Shaanxi Provincial Social Science Fund (2018M37) and Shaanxi 13<sup>th</sup> Five-year Plan Teaching Research Program (SGH18H027).

## References

- [1] Bortfeld, H. et al. Disfluency rates in spontaneous speech: Effects of age, relationship, topic, role, and gender [J]. *Language and Speech*, 2001, (44): 123-147.
- [2] Allen, V. 1971. Teaching intonation, from theory to practice [J]. *TESOL Quarterly* (4): 73-91.
- [3] Clark, J. & C. Yallop. 1995. *An Introduction to Phonetics and Phonology* [M]. Oxford: Blackwell.
- [4] Halliday, M. A. K. 1967. *Intonation and Grammar in British English* [M]. Berli: Mouton.
- [5] Halliday, M. A. K. 1970. *A Course in Spoken English* [M]. London: Oxford University Press.
- [6] Anderman, G. M. & Rogers, M. (eds.) (2008). *Incorporating corpora. The linguist and the Translator*. Clevedon, Multilingual Matters.
- [7] Liu Miqing *Theoretical Study on Translation and Interpreting* [M], Beijing, China Translation and Publishing Corporation, 2006.
- [8] Wen Qiufang, Wang Jinquan *Parallel Corpus of Chinese EFL learners* [M]. Beijing, FLTRP, 2008.
- [9] Yang Jun, *Research on Disfluent Spoken Utterances* [J] *Foreign Language Teaching and Research*, 2004, (4): 87-91.
- [10] Hu Kaibao, Tao Qin. (2010). Establishment and Application of C-E Conference Interpreting Corpus. *Chinese Translators Journal* (5), 49-56.
- [11] Chen Hua (2006) *English Intonation Pattern and Its Acoustic Realization*, *Foreign Language Study* (99), 9-18.
- [12] K. Plevoets, B. Defrancq 2016, The effect of informational load on disfluencies in interpreting. A corpus-based regression analysis, *Transl. Interpret. Stud.*, 11 (2) (2016).
- [13] P. Mead (2005) Methodological issues in the study of interpreters' fluency *Interpreters' Newslett.*, 13 (2005), pp. 39-63.
- [14] Y Jiang (2020). Effect of dependency distance of source text on disfluencies in interpreting, *Lingua* (243), 102873.
- [15] Y. Lin, Q. Lv, J. Liang, 2018, Predicting fluency with language proficiency, working memory, and directionality in simultaneous interpreting *Front. Psychol.*, 9 (2018), p. 1543.