**What I have done:**

1. Booked the Corpus Linguistics summer school of UCL ( payment made), then organized and submitted all the documents to apply for Labex EFL funding.
2. Read further more papers (in Chinese) by the presenter ([Zhongmin Chen](https://imoll.fudan.edu.cn/info/1017/1025.htm)) in conjunction with the lecture I took last week ( Sound cues and their evolution):

√Motor theory and language cognition (2015)

√On the Interactive Mechanism Between Speech Pronunciation and Perception (2019)

√Retroflexed sounds (2022)

√Anatomical and physiological mechanisms of tone production (2022)

√The reasons of sound change (2022)

√The Characteristics of Speech Perception and Their Anatomical and Physiological Mechanisms (2022)

Actually, I gained a lot from these.

Since I got the research topic of reduction last March and started to look up & read the literature, I kept accumulating questions, solving problems, and discovering new ones ....... In that immature research proposal submitted in May, I tried to give a brief overview of speech reductions research in terms of mechanisms, influencing factors, and cross-linguistic performance. At that time I gathered a lot of abstract knowledge or conclusions (I knew, I remembered, and I understood, but I didn't really realize what those conclusions really meant when integrated). However, by then I had a vague sense that if I wanted to improve the robustness of automatic speech recognition (I read a lot of papers related to this part earlier this year and got a better understanding of it), I had to do a deeper study of the mechanisms and concrete manifestations of reduction. So how to study it? I proposed that we should start from the stability features of a particular segment that can still be preserved in various contexts, while those features that are extremely vulnerable to changes in the surrounding contexts are not necessary for us to identify it, and it is precisely these features that cause the reducing variation of speech. But my theoretical knowledge base was not sufficient and my explanation of this idea was shallow, so I put it aside for a while. But recently reading these papers did inspire me with new ideas about this research line.

First, we need to acknowledge that reduction is itself a phonetic variation, and it is the variation that occurs most frequently in everyday communication. Structuralist linguistics divides the study of language into two dimensions, diachronic and synchronic, and I think this division is very meaningful. Phonetic variation occurs at every historical stage, and it is the synchronic variation that eventually leads to the diachronic change, because as the high-frequency variation gradually spreads out in a particular speech community and is preserved as a stable form in their mental lexicon, then this new form replaces the prototype that it was originally stored in the human brain, resulting in a lexical update. If the result of this update can be passed on to the next generation, then it naturally leads to the sound changes. We can find many examples of such changes in English, French, and Chinese. So the first concept I want to explain here is the distinction between "sound variation" and "sound change". The former is the variation in the speech flow at the synchronic level, while the latter is the longitudinal change that occurs from ancient to modern times after the triumph of variation, the latter being the origin of the former. This view coincides with the claim of John Ohala (1989).

Second, whether it is sound change or sound variation, they both have a certain pattern and rules, otherwise our language will become extremely confusing and even affect communication. But from Neogrammarians to structuralist linguistics, the Prague school, and even Chomsky's transformational-generative theory, they have been more concerned with sound change over the past 200 years, and with sound variation, they either think it is impossible to study, or that it is not worth studying (Chomsky thinks it is language performance, which varies so much that he is concerned with the process of differentiation of language performance from universal human language ability, in other words, the mechanism of derivation from the underlying form to the surface form). This situation is largely related to the long affiliation of modern linguistics with literature before it became independent. Both Western translations of religious literature and Eastern interpretations of ancient Chinese texts, had to explain the fundamental question of "why was the language different from the present". Therefore, sound change received a great deal of attention from early linguists. It was only with the introduction of sociolinguistics (Labov, Weinreich & Herzogo, 1968) and lexical diffusion theory (William S-Y. Wang., 1969) that the synchronic sound variation finally caught the eye of linguists. Sociolinguistics views language as an ordered heterogeneous entity, and language changes only in the population, so that the social properties of the people themselves must be taken into account in language variation. The ultimate goal of sociolinguistics is to "find the leaders of change," and by analyzing the linguistic characteristics of people with different social attributes, they model the chain of language transmission among communities. The lexical diffusion theory is also concerned with synchronic sound variation and suggests, for the first time, that we should distinguish between two different stages of sound variation: actuation and implementation. The former is the trigger of sound variation, such as the speaker's pronunciation, the transmission of sound waves, and the listener's perception on it. The latter refers to the diffusion of sound variation in the community, which is similar to the research paradigm of sociolinguistics. In our study, the focus of attention is on the actuation of speech reduction.

Another central claim of the lexical diffusion theory is that sudden sound changes v.s. gradual lexical changes. Unfortunately, because the technique of phonetic research was not abundant at that time, it was difficult for linguists to concretely observe the changes in the internal details of phonetic features, so they localized the smallest unit of sound change to the phonetic segment (at most the phonemic features of the phonological system), thus concluding the hypothesis of " sudden phonetic change ". However, just as in chemistry, where the study of particles of different sizes (molecules, atoms, neutrons, protons, etc.) is advancing, the study of the various variations of speech can be greatly improved when the smallest unit of speech is no longer a segment, but a more specific phonetic cues. Sound is not a suddern change; it is equally gradual.

We can take a metaphor. We see a bee flying in a blossom, and we might say generally, "Look! There's a bee flying." This is when we look at the bee whole as a unit. But is the bee really flying all over? No, it is the bee's vibrating wings and legs that play a key role in maintaining balance, so we can see that the bee is moving and flying steadily. Similarly, it is not rigorous enough to say in general that the phonetic segment changed or did not change, and our main concern now is how to push the study to the fine detail. If the sound has changed, what specific articulatory, acoustic, and perceptual features have changed? The changes in these features can be visually observed and analyzed with the help of rich and advanced technological tools. These multi-dimensional features that affect speech changes are collectively referred to as sound cues.

Do all the sound cues of a single phone in a context necessarily occur? The answer is still no. Some cues can be expressed in the segment itself, such as the formant of vowels, etc. We refer to these cues as internal cues, while some cues can only be expressed when they are near the adjacent segment, and when the adjacent segment is lost, these cues become weak cues and are easily transformed, such as the track between vowels and consonants, VOT, etc. We refer to these cues as external cues. Also, the expression of segmental features is not one-dimensional and linear like text transcription, but partially overlapping; if the individual cues of two segments are in conflict, listeners will pay more attention to those strong cues.(There is an interesting little experiment in this section, which I will present in next Wednesday's report)

External cues and internal cues have sometimes transformed into each other throughout history, and their status has not remained static. But whenever this happens, it must mean that the landscape of the language has changed dramatically. The creation of Chinese tones is a typical example. In the earlier stage, Chinese had a voiced/voiceless opposition of plosives without any tones. Thus, the pairing "voiced plosives + vowel" evolves into a low tone, and the pairing "voiceless plosives + vowel" evolves into a high tone. In this case, the various effects of the voiced/voiceless plosives on the subsequent vowels, as external cues, eventually led to the loss of not only the voiced/voiceless opposition, but also to the transfer of this opposition to the new suprasegmental opposition, which led to the creation of tones, thus significantly changing the structure of the Chinese language. (The creation of Chinese tones is very complicated. From the phonological characteristics of minority languages in China, it is possible that Chinese has undergone a evolutionary path of "voiced/voiceless→breathy/modal→low tone/high tone", which means from the opposition of voiced/voiceless plosives to the opposition of high/low tones, the breathy/modal phonation type opposition is also an important stage. The cue of the breathy voice often occurring when the vowel is followed by the voiced plosives may also provide a valid explanation for this evolutionary path.)

The division between external cues and internal cues was proposed by Zhongmin Chen in 2022, and this theory is still focused on explaining sound change. However, based on the previous description, I think it is also very enlightening for the study of reduction, and this theory adds more specificity and credibility to the idea that I proposed last year. In synchronic sound variation, especially by comparing reduced forms of the same sound in different contexts, we can identify and test which cues are internal and which are external, with external cues being easily reduced in the speech flow and internal cues not. Completing this work not only can integrate the reduced features from refinement studies into speech models to better describe the real speech signal, enhance the generalization ability and adaptability of speech models, and thus improve the accuracy of automatic speech recognition, but also can expand our knowledge of speech variation at the synchronic level, which is also significant for the theoretical study of speech ontology.

After all, the continuous analysis of sound cues smaller than phonetic segment indicates in itself the deepening of phonetic research.

**What I want to do:**

1. There are so many sound and sound cues to dig out and I really need to be careful and decide as soon as possible which areas I want to focus.
2. Follow the new manual instructions sent by Philipp to practice using the Montreal Forced Aligner.