Aspirated stops: situation in Mandarin, English, and Czech

Hana TŠsková, Ph.D. (廖敏), triskova@orient.cas.cz
Oriental Institute of the Czech Academy of Sciences, Prague

Paper presented at the 6th conference of the Phonetic Association of China,
Beijing (IL CASS), May 21, 2011

Introduction

Unlike the rest of the participants of this conference, I am not a phonetician. I do not pursue my own experimental phonetic research. I am a Sinologist interested in phonetics. The main goal of my work is to introduce the results of phonetic research acchieved by others into the process of teaching Chinese as a second language. I believe such efforts may considerably improve the traditional ways of teaching Mandarin pronunciation and make it more efficient.

Aspiration

Aspirated consonants (mainly stops, less often affricates) exist in many languages of the world. As far as a **definition of aspiration** is concerned, there is not an agreement. Let me quote one of them (Ladefoged & Maddieson 1996:70): ñ*Aspiration is a period after the release of a stricture and before the start of a regular voicing*é *in which the vocal folds are markedly further appart than they are in modally voiced sounds.***ñ** In this definition, a major definitory feature of an aspirated consonant is a **markedly positive VOT**. Yet there are various other definitions of aspiration. It is probably best to view aspiration as a **complex of several articulatory and acoustic features**, not as a single feature. This is the view of Wu Zongji (2004) and also of Zhu Chuan (1997:65) who writes: 现代语音研究表明，送气和不送气音不是单一特征上的对立。两者的区别有：气流的强弱，气压的大小，VOT 值的长短，声源特征的差异等。

Voice Onset Time

Just to make brief overview, I am giving here a sketch of three major situations as far as VOT is concerned:
- In the **voiced** stops, VOT is **negative.** An example is a fully voiced Czech stop **[b]** in the words ***b****ota*, ***b****azén*, ***b****udka*.
- In the **voiceless** stops, VOT is **about zero**. An example is a voiceless stop **[p]** in Mandarin words ***b****Ūng* 帮, ***b****ù* 不, ***b****ái* 白, or in Czech words ***p****ísek*, ***p****án*.
- In the **aspirated** stops, VOT is **positive**. An example is a Mandarin aspirated stop **[p^{h}]** in Mandarin words *pàng* 胖, *pi****Ū****o* 漂, or in English words ***p****encil*, ***p****eace*, ***p****aper*.

(1)	**voiced [b]**		negative VOT
(2)	**voiceless [p]**		zero VOT
(3)	**aspirated [p^h]**		positive VOT

Of course, there can be many different situations in coordination between the opening of a closure and the activity of the vocal folds. In other words, **VOT is a continuum**. At one extreme, there are **fully voiced stops** with a markedly negative VOT. The examples are the Czech stops **[b], [d], [g]** (e. g. in the word ***b****ota, **d**ali, **g**uma*). They may have about the following VOT values (my pronunciation): **[b] =** minus 143ms, **[d]** = minus 108ms, **[g]** = minus 138ms. Then there are **partly voiced stops** with only a slightly negative VOT. For instance English initial **[b], [d], [g]**. Further along the continuum, there are **voiceless stops**. They have about a zero or slightly positive VOT. The examples are Czech stops **[p], [t], [k]** (e.g. in the words ***p****ísek, také, **k**olik*)**.** They may have the following VOT values (my pronunciation): **[p]** is 2ms, **[t]** is 13ms, **[k]** is 26ms. Further on we can find **weakly aspirated stops** such as English intial **[k^h]** (***k****eyboard*) which has VOT about **80 ms**. Finally, at the other extreme of the continuum, there are **strongly aspirated stops** with a large positive VOT. An example is **[k^h]** in Navajo: its VOT is **160 ms** which is twice than in English **[k^h]**.

Laryngeal features

In phonology, the activity of the vocal folds can be expressed by means of distinctive features **[±voice]** and **[±aspirated]**. They fall into a group of so called **laryngeal features**. Let us see **how the laryngeal features [±voice]** and **[±aspirated]** can be used for **making phonological distiction** in languages. Theoretically, there are four possible categories of consonants distinguished by various combinations of ñ+ò and ñï ñ values:

ï <u>voiceless unaspirated</u>	**[ï voice] [ï aspirated]**
ï voiceless aspirated	**[ï voice] [+aspirated]**
ï voiced unaspirated	**[+voice] [ï aspirated]**
ï (voiced aspirated)	**[+voice] [+aspirated]**

The first, underlined category is probably used **almost universally** across the languages. On the other hand, the fourth category is **quite rare** (it is found e.g. in some Indian language such as Hindi, which has voiced aspirates: **/b^h/, /d^h/é**). Let us leave this unusual category aside and assume that, for making contrast, the languages choose from **the first three categories**. They may pick up either **two**, or **three** categories.

Languages contrasting two categories

Most languages pick up two categories. E.g. Czech, French, Russian, Spanish, Italian, Mandarin, English and many others. The first category ï **voiceless unaspirated consonants** ï is basically always used. The second category added to it to make contrast is either the **voiceless aspirated consonants**, or the **voiced unaspirated consonants.** That means the language accepts either an **aspiration contrast**, or a **voicing contrast.**

a) **aspiration contrast:** voiceless unaspirated X voiceless aspirated
b) **voicing contrast:** voiceless unaspirated X voiced unaspirated

Voicing contrast is more common in the languages of the world.

Languages contrasting three categories

Less often a **three-sided contrast** can be found in languages (that means all three categories are used). The examples are: Armenian, Thai, Burmese, or Shanghainese. For instance in Shanghainese there is a contrast between **/b/, /p/**, **/p^{h}/**: 抱 **[bɔ]** ñto hugñ, 报 **[pɔ]** ñnewspapersñ, 泡 **[p^{h}ɔ]** ñto soakñ (without tones).

We shall be concerned with **Mandarin, English and Czech** ï all of them use a two-sided contrast.

Aspirates in Mandarin

Mandarin uses a **contrast of aspiration**. That is, aspiration is used as a **distinctive feature**. If a speaker fails to add aspiration, it may lead to confusion. A popular example is: *Dùzi b**ʁ**o le* 肚子饱了 vs. *Tùzi p**ʁ**o le* 兔子跑了. This goes to say that failing to add aspiration is not just <u>some</u> mistake, but a <u>serious</u> mistake. Mandarin has six aspirates: three stops (*p, t, k* in *pinyin*) and three affricates (*c, ch, q* in *pinyin*). Note that aspirated affricates are rather rare cross-language.

Aspirates in English

The situation in English is different than in Mandarin. The crucial distinctive feature in the system of consonants is **voicing**. There are three pairs of stops, where the first member is **voiceless**, and the second member is **voiced**: **/p/ ï /b/**, **/t/ ï /d/**, **/k/ ï /g/**. However, the aspirated consonants do occur in English phonetically: if I pronounce the words ***p**in, **t**oo, **k**ey*, or ***p**encil, **t**ime, **k**eyboard*, you can hear that the voiceless phonemes /**p, t, k/** carry certain amount of aspiration. The aspiration basically emerges only if **p, t, k** are **in the initial position**. Another context supporting the emergence of aspiration is the beginning of a **stressed syllable** ï it concerns also the stressed syllables <u>within</u> the words (*po**t**ato*). In other situations, the aspiration of /**p, t, k/** is either rather light, or nonexisting. The situations without aspiration are the

following. **First**, if /**p, t, k/** occuring in the onset of an initial syllable of a word is preceded by **/s/** (*s**p**in, s**t**uck, s**k**ies*). **Second**, if the initial syllable is unstressed (***p**olite*). **Third**, at the end of a word (*li**p**, bi**t**, stic**k***).

So in English, aspiration exist only as an **accompanying phonetic feature** of the stops /**p, t, k/ in certain positions**. It is fully predictable according to the position of a stop. The aspirated consonants do not have a status of phonemes, aspiration does not distinguish the meanings of words. Thus, neglecting aspiration does not seem to be a serious mistake. For instance the word *pencil* should be correctly pronounced as **[phensl]**. If I pronounce it as **[pensl]**, a native speaker of English will know what I wanted to say ï although he will immediately recognize I am not a native speaker. However, forgetting to use aspiration in pronunciation of the English initial stops /**p, t, k/** can cause confusion in certain contexts. We shall come back to this topic later.

Aspirates in Czech

The situation in Czech is very simple ï aspirated consonants **do not exist there** ï neither **phonologically** (as Czech uses a **voicing contrast**), nor **phonetically:**

- Czech **voiceless stops** have about zero VOT or very mildly positive VOT. They can never reach the VOT values of the English initial stops **[p^h], [t^h], [k^h]**.
- Czech **voiced stops [b], [d], [g]** etc. have a markedly negative VOT, as mentioned before.

The Czech speakers usually have a very vague or no idea about what aspiration is. They are not sensitive to it.

Phonetic properties vs. phonological categories

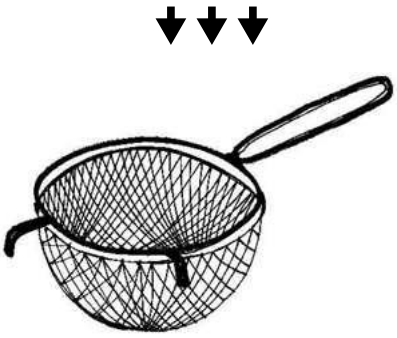
Let us go back to English now. A traditional analysis of the English phonological system of consonants accepts a **voicing contrast**, as mentioned before. The stops /**p**, **t**, **k/** are [ï *voice*], while the stops /**b**, **d**, **g/** are [+*voice*]. However, **in perception** it is actually **aspiration**, **not voicing** which distinguishes the initial /**p, t, k/** from the initial /**b, d, g/**. If we pronounce the stops /**b**, **d**, **g/ without any voicing**, a native speaker will still perceive them as voiced phonemes /**b**, **d**, **g/**. For instance, if I say ***bye bye*** and pronounce the initial **/b/** as **completely devoiced [b̥]**, a native speaker will still correctly perceive this sound as a **voiced phoneme /b/**. The reason is he did not hear any aspiration, which is the major perceptional cue for the unaspirated stop **/p/**. Indeed, the English initial /**b**, **d**, **g/** are frequently realized as lax, devoiced **[b̥]**, **[d̥]**, **[g̊]**. **Fully voiced stops [b]**, **[d]**, **[g]** may occur only between two vowels (*obey*, *audience*). Now, what happens if a native speaker hears fully voiced **[b]**, **[d]**, **[g]** with a strongly negative VOT in the initial position? He will correctly identify them as the **voiced phonemes /b, d, g/**, but they will sound quite strange to him (by the way, this is a common mistake of Czech speakers of English). On the other hand, if I pronounce

pie without any aspiration, that is as **[paːɪ]** instead of **[p^haːɪ]**, a native speaker will hear it as ***bye***.

We found out that in English it is not voicing what distinguishes the initial voiceless stops **/p, t, k/** from the voiced stops **/b, d, g/**. How come then that voicing is accepted as a distinctive feature? Some authors suggest another solution: they refuse the **contrast of voicing**, and accept the **contrast tenseï lax**. Instead of [±*voice*], they establish a distinctive feature [±*tense*]. The stops **/p, t, k/** are [+*tense*], the stops **/b, d, g/** are [ï *tense*].

As illustrated before, the **aspirated stops [p^h]**, **[t^h]**, **[k^h]** have a **different phonological** status in English and in Mandarin: in **Mandarin** they are indeed aspirated phonemes **/p^h/, /t^h/, /k^h/** etc. On the other hand, in **English** they are only positional variants of unaspirated phonemes **/p/, /t/, /k/**. Further, we observed differences in pronunciation of the **voiceless phonemes /p, t, k/** in English and in Czech: in English they must be aspirated in the initial position, while in Czech they never become aspirated. This all looks quite confusing. It seems it is not possible to define what is a ñvoiceless consonantñ, or a ñvoiced consonantñ, or an ñaspirated consonantñ for all languages universally. It is indeed so. The identical physical sounds may be evaluated as different sorts of phonemes (ñvoiceless ñ, or ñvoicedñ, or ñaspiratedñ) by native speakers of different languages. E.g. a labial voiceless stop **[p]** in an initial position will be evaluated as a voiceless phoneme **/p/** by a Czech native speaker, but as a voiced phoneme **/b/** by a native speaker of English. The native speakers of different languages have different **ñphonological sievesñ** through which they drive the perceived physical sounds:

[p^h] [b] [k^h]
[g] [d] [t^h] [k]
[ɡ̊] [t]



/p/ /g/ /k/
/d/ /t/
/b/

Let me sum up the situation in **English, Czech and Mandarin** for the initial stops.

	voiced stop phonemes /b, d, g/	**voiceless stop phonemes /p, t, k/**	**aspirated stop phonemes /p^h, t^h, k^h/**
pronunciation in English	**[b̥], [d̥], [g̊] zero VOT, lax**	**[p^h], [t^h], [k^h] positive VOT, tense**	- - -
pronunciation in Czech	**[b], [d], [g] negative VOT, lax**	**[p], [t], [k] zero VOT, tense**	- - -
pronunciation in Mandarin	- - -	**[b̥], [d̥], [g̊] zero VOT, lax**	**[p^h], [t^h], [k^h] positive VOT, slightly more tense**

Commenting on the table:

We observe that the phonemes **/b, d, g/, /p, t, k/**, and **/p^h, t^h, k^h/** (initial position) have different phonetic properties while pronounced in the examined three languages:

ŏ The first column ï the sounds perceived as the **voiced phonemes /b, d, g/:**
In English they have about a zero VOT. They are lax. They can be transcibed as devoiced **[b̥], [d̥], [g̊]**.
In Czech they typically have a **large negative VOT**.
In Mandarin they do not exist.

ŏ The second column ï the sounds perceived as the **voiceless phonemes /p, t, k/:**
in English they have a **positive VOT** (i.e. they are aspirated). They are **tense**. They can be transcribed as **[p^h], [t^h], [k^h]**.
In Czech, they have about a **zero VOT** (i.e. they are voiceless). They are rather **tense**. They can be transcribed as **[p], [t], [k]**.
In Mandarin, they also have a **zero VOT** (i.e. they are voiceless), but they are **lax**. They can be transcribed as devoiced **[b̥], [d̥], [g̊]**.

ŏ The third column ï the sounds perceived as the **aspirated phonemes /p^h, t^h, k^h/:**
In Mandarin, they have a markedly positive **VOT**. They are somewhat tense. They can be transcribed as **[p^h], [t^h], [k^h]**.
In English and Czech, they do not exist.

As we can conclude, **it is really impossible to define** a ñvoiced consonantñ or ñvoiceless consonantñ or an ñaspirated consonantñ for all languages in the same way. That means they cannot be defined in terms of some **constant ranges of VOT values** applying to all alnguages. The phonetic cues which decide about perception of voicing contrast and aspiration contrast may be different in different languages.

Polarization principle

As pointed out before, there are three main categories at disposal to choose from for making phonological contrast: **voiced, vociceless unaspirated, and aspirated**. We observed that the way the particular categories may be phonetically realized can be variable. This variability may be explained by a ñ**polarization principleò** (Patricia Keating 1984). She claims that if a language chooses **two categories**, there is a tendency **to enhance the phonetic contrast** between them.

	voiced stops /b, d, g/	**voiceless stops /p, t, k/**	**aspirated stops /p^h, t^h, k^h/**
English	**[b̥], [d̥], [g̊]** **zero VOT**	**[p^h], [t^h], [k^h]** **weakly positive VOT** (i.e. adding slight aspiration) →→→	- - -
Czech	**[b], [d], [g]** **large negative VOT** ←←←	**zero VOT** **[p], [t], [k]**	- - -
Mandarin	- - -	**[b̥], [d̥], [g̊]** **zero VOT**, but lax and possibly sonorized ←←←	**[p^h], [t^h], [k^h]** **positive VOT**

Commenting on the table:

Let us suppose that all languages have a category of **voiceless unaspirated consonants** According to the ñpolarization principleò, they may be realized in **two different ways** ï depending on what is the other category accepted for making a contrast:

a) If a language accepts an **aspiration contrast**, then the category of the voiceless unaspirated consonants will probably tend to be somewhat **sonorized** (that is with a small negative VOT). This is the case of **Mandarin**: the voiceless phonemes (be it stops, or affricates) are realized as **lax** (which is normally typical for the voiced consonants); furthermore, they may become more or less **voiced** in some conditions (especially in unstressed syllables).

b) If a language accepts a **voicing contrast**, the category of voiceless unaspirated consonants will be probably somewhat **aspirated**. This is the case of **English**. However, this is not the case of **Czech** ï here the contrast is enhanced by other means: a large negative VOT in the voiced stops.

Foreign accents

Let me sum up the situation in the three languages that have been examined ï Mandarin, English and Czech

ð First, **Mandarin** uses the **aspiration contrast**.

ð Second, **English and Czech** both use **voicing contrast**. However, they use different VOT values for realization of the same phonological categories: **voiceless phonemes /p, t, k/** are pronounced with a positive VOT in English, while with about a zero VOT in Czech; **voiced phonemes /b, d, g/** are pronounced with about a zero VOT in English, while they have a markedly negative VOT in Czech.

These facts are closely related to different **ñphonological sievesñ** of the native speakers of English, Mandarin and Czech. In other words, they evaluate the identical physical sounds in different ways. The differences in speaking habits and phonological sieves contribute to various **typical mistakes** or accents:

ï **Native speakers of English speaking Czech:** undue aspiration of the Czech initial stops **/p**, **t**, **k/**: they pronounce them as **[p^{h}]**, **[t^{h}]**, **[k^{h}]** (e.g. *také* **[takeː]** ñalsoñ is pronounced as ***[t^{h}akeː])**. As aspiration does not function as a distinctive feature in Czech, the mistake is not overly serious: the result is òonlyñ an impression of a weird pronunciation for the ears of a Czech listener.

ï **Native speakers of Czech speaking English:** failing to add aspiration to the initial stops **/p**, **t**, **k/,** pronouncing them as **[p]**, **[t]**, **[k]** instead of **[p^{h}]**, **[t^{h}]**, **[k^{h}]** (e.g. ***t**oo* is wrongly pronounced as ***[tuː]**). This mistake is slightly more serious. It is true that aspiration does <u>not</u> function as a distinctive feature in English. However, it is still an important perceptional cue for the initial /**p, t, k**/. Another typical Czech mistake is pronouncing English **/b, d, g/** as fully voiced **[b]**, **[d]**, **[g]** with a large negative VOT (e.g. in *day* **[deɪ]**).

ï Native speakers of English speaking Mandarin: they often have problems with the Mandarin voiceless unaspirated stops **/p, t, k**/ (*b*, *d*, *g* in *pinyin*). As Kratochvíl 1968:26 remarks, they tend to add either aspiration, or voicing. E.g. in the word ***bīng*** 兵 [**piŋ**]. The beginners often pronounce the initial as fully voiced: **[biŋ]**. After some time of learning they are advanced enough to know there <u>must not</u> be any voicing for Mandarin voiceless stops **/p, t, k**/ (*b*, *d*, *g*). While attempting to get rid of voicing, they may end up adding aspiration ï saying **[p^{h}iŋ]**. The reason is that, in their speaking habits, the initial voiceless phonemes /**p, t, k**/ must carry aspiration in pronunciation. Technically speaking, they carry the increasing of the VOT value too far, making it positive instead of just about zero. Such mistake is serious and may lead to confusion.

ï **Native speakers of Czech speaking Mandarin:** The beginners learning new words may have problems to remember a distinction between the aspirated **/p^{h}**, **t^{h}**, **k^{h}/** (*p*, *t*, *k* in *pinyin*) and unaspirated **/p, t, k**/ (*b*, *d*, *g* in *pinyin*). For instance they confuse *kuang* and *guang*. In speech production, however, they usually do not have problems. Occassionally they may exaggerate the **[h]** character of aspiration, pronouncing it too strongly as a separate sound. Also they may misplace the source of aspiration friction. E.g. they may pronounce *qì* 气 as **[tʃ^{h}iː]** instead of correct **[tɕ^{ɕ}iː]**.

ï **Native speakers of Mandarin speaking Czech:** they regularly have problems to distinguish between the **voiceless stops /p, t, k**/, and the **voiced stops /b, d, g**/ ï both in speech perception and in speech production. They just do not hear the difference. They can for instance confuse the Czech words ***pár*** (ña coupleò) and ***bar*** (酒吧 *jiœba*) (note that, unlike in Mandarin, vowel quantity is distinctive in Czech). Another example: ***tÍo*** (ña bodyò) and ***dÍo*** (ñcanoonò).

Conclusion

Foreign students of Mandarin come from many different countries. They are native speakers of a wide range of languages. Their languages choose **various categories** for making phonological contrast among the stop consonants. These categories may be realized with various ranges of VOTs. Consequently, the **phonetic phenomena of aspiration and voicing** may be perceived and phonologically interpreted in different ways by different students. Further, while speaking, the students have different habits, being used to employ different ranges of VOTs for particular stop consonants. All this results in **different sorts of mistakes** they make ï both in speech perception and in speech production. To correct such mistakes in language teaching it is usefull to know their roots.

LITERATURE

CRUTTENDEN, Allan ed. *Gimson´s Pronunciation of English.* London: Arnold, 2001.

CRYSTAL, David. *Dictionary of Linguistics and Phonetics.* Oxford: Blackwell Publishers, 1997.

GIEGERICH, Heinz. *English Phonology. An Introduction.* Cambridge: Cambridge University Press, 1992.

GUSSENHOVEN, Carlos ï JACOBS, Heike. *Understanding Phonology.* Beijing: Waiyu jiaoxue yu yanjiu chubanshe (in cooperation with Blackwell Publishers), 2001.

KEATING, Patricia. Phonetic and Phonological Representation of Stop Consonant Voicing. *Language*, 1984, No. 2, pp. 286-319.

KRATOCHVÍL, Paul. *The Chinese Language Today.* London: Hutchinson University Library, 1968.

LADEFOGED, Peter ï MADDIESON, Ian. *The Sounds of the World´s Languages.* Oxford: Blackwell Publishers, 1996.

LADEFOGED, Peter. *A Course in Phonetics.* Boston: Thomson Wadsworth, 2006.

LIN, Tao ï WANG, Lijia. *Yuyinxue jiaocheng*. Beijing: Beijing daxue, 2003.

LIN, Yen-Hwei. *The Sounds of Chinese.* Cambridge: Cambridge University Press, 2007.

LISKER, Leigh ï ABRAMSON, Arthur. A Cross-language Study of Voicing in Initial Stops: Acoustical Measurements. *Word*, 1964, No. 3, pp. 384-422.

MADDIESON, Ian. *Patterns of Sounds.* Cambridge: Cambridge University Press, 1984.

ROACH, Peter. *English Phonetics and Phonology*. Cambridge: Cambridge University Press, 1996.

WU, Zongji. Putonghua fuyin bu songqi ï songqi qubie de shiyan yanjiu. In: Wu, Zongji. *Wu Zongji yuyanxue lunwen ji.* Beijing: Shangwu yinshuguan, 2004. pp. 31-65.

WU, Zongji et al. *Xiandai hanyu yuyin gaiyao*. Beijing: Huayu jiaoxue chubanshe, 1992.

ZHU, Chuan ed. *Waiguo xuesheng hanyu yuyin xuexi duice*. Beijing: Yuwen chubanshe, 1997.