ARIPUC 20, 1986, p. 79-98

CAN "THE GREAT TONE SPLIT" IN THAI BE PHONETICALLY EXPLAINED?*)

JØRGEN RISCHEL

This paper deals with the so-called "Great Tone Split" in Thai, which in most dialects has taken a course contrary to the general contention that initial voiced and voiceless consonants, if they have any conditioning effect on tonogenesis, will trigger a difference between a lower tone in the former case, and a higher tone in the latter case. On this background it is discussed to what extent tonogenesis of this type is at all phonetically explicable at present.

I. INTRODUCTION

Explanation of sound change is one of the major tasks of current theoretical work, although researchers do not agree on how to deal with this issue. Among the types of change which seemingly have been explained successfully in recent years, a certain type of "tonogenesis" occupies a prominent place. Does it so rightfully?

The tone systems of numerous Southeast Asian languages are known to have developed out of an earlier pattern of syllable-final and syllable-initial features in the language in question.

---

*) *A substantial part of this paper was presented under the title "Tonogenesis in Thai: A Phonetic Paradox", at the Sino-Tibetan Conference, which took place in Bangkok in August 1985. For a more general presentation of some of the controversial features of tonogenesis in Thai, see Rischel (1985), especially p. 64-68. The reader is referred to that source for further references to the relevant literature on Thai (Bibliography loc.cit. p. 81-93).*

This obtains not only for Sino-Tibetan but also for Austro-asiatic and Tai languages (as shown by the work of Haudricourt and others), i.e., it is an areal feature. What has most generally attracted the attention of phonologists interested in "tonogenesis" (to use the excellent term coined by Matisoff) is the role played by initial consonants via the different (more or less automatic) pitch perturbations they induce on vowels, depending on their inherent manner features such as voicing, aspiration, and glottalization.

In discussing how such pitch perturbation by initials can give rise to tones one must constantly pay attention to the remaining, relevant properties of the syllables. Phonologization of initial pitch perturbation may well make a non-tonal language tonal (Modern Tibetan is interpretable as a case in point), but often the result is a tone split within an already existing tonal system, as is indeed the case in many East Asian languages and dialects. Obviously, sound change has both an articulatory, a perceptual, and a high-level organizational aspect. Therefore, one cannot hope to arrive at a unified explanation of any mechanism of tonal change without taking into consideration the psychological role played by already existing or incipient tonal distinctions (e.g. such that have to do with syllable quantity or types of syllable termination), and the degrees of freedom inherent in the tonal development. Thus, the conditions for tonogenesis must be considered anew in all detail for each particular language, and with reference to all chronological stages that seem relevant to the explanation.

Research on Tai languages and on the dialects of Thai proper within the last quarter of a century has shown that the interplay among such factors conditioning tonal development may lead to widely different tone patterns for closely related languages or dialects. The possibility of nonetheless finding general principles of development for these languages and dialects has been the object of important recent work (Egerod 1971, Brown 1975, Strecker 1979).

As for the role of initial consonant type, the by now trivial phonetic observation is that after a voiced initial (at least after a voiced stop) the pitch on the following vowel tends to start relatively lower than after a voiceless initial. It is, however, important to bear in mind that the degree of (negative or positive) pitch perturbation may depend on very specific articulatory properties of the consonants (including such that are not always indicated in phonemic descriptions of languages), so that one cannot expect to be able to predict the degree of pitch perturbation simply from a knowledge of the voicing state of the consonant. Thus, it is a controversial issue how the series of Tai consonants reconstructed by Li (1943) as glottalized voiced stops might be expected to influence the pitch (this particular issue has been discussed at length by Donna Erickson (1975) with reference to tonogenesis in Thai). More generally, it must be stated that there is not yet much of a general phonetic theory explaining why voiced and voiceless initials condition a pitch difference in the following segment.

Hombert, Ohala and Ewan (1979, p. 42-45) evaluate the relative merits of three types of phonetic explanation, viz. the Aerodynamic Hypothesis, the Halle-Stevens type of Vocal-Cord Tension Hypothesis, and the Vertical Tension Hypothesis: (1) The aerodynamic hypothesis is based on the observation that the initial airflow in voiced stops just after release is low compared to that of voiceless (particularly aspirated) stops, which would make the vocal cords vibrate more slowly just after release in the former than in the latter case. As pointed out by Hombert et al., this fails to explain why the pitch perturbation is normally seen to extend far into the vowel even though the aerodynamic effect should disappear shortly after the release. - (2) The first tension hypothesis takes its point of departure in Halle and Stevens' suggestion that there is a difference in horizontal vocal cord tension, voiced stops being articulated with slack but voiceless stops with stiff vocal cords, and that these states affect the Fo of adjacent vowels. This should be true preconsonantally as well as postconsonantally, which is definitely not corroborated by empirical evidence (Kohler 1985, p. 22 suggests that the similar effect of French unaspirated and English aspirated voiceless stops may be due to a shared feature of tension in initial position, however). - (3) The second tension hypothesis suggests that Fo perturbation has to do with a difference in vertical tension of the vocal cords. This is apparently corroborated by the observation that the larynx position is lowered for voiced as opposed to voiceless stops and that this difference in larynx position persists well into the following vowel. Unfortunately, Hombert et al. state: "the voiced stops' effect on Fo is like that of the sonorants, and it is the Fo after the voiceless stops which is perturbed, i.e. raised above them" (p. 45). Nevertheless, the authors lean to this hypothesis with a view to the difficulties facing the aerodynamic hypothesis.

Recent research on larynx height (Riordan 1980, Reinholt Petersen 1983) does not, however, support the assumption that there is a direct and simple relationship between larynx height and Fo perturbation in vowels after obstruents.

Kohler (1985, with references also to recent work by Haggard and others), on the contrary, finds more support for the aerodynamic hypothesis, possibly coupled with horizontal tension differences. He points out that most earlier data muddle the issue by mixing up sentence intonation contour with local pitch perturbation. - At present it seems safest to conclude that although there is obviously a principled connection between voicing state and Fo, the mechanism involved is not (at least not fully) understood. In fact, it is still an open question to what extent such pitch perturbation is at all an automatic phenomenon (whatever its mechanical cause), and to what extent it is actively programmed as a more or less independent variable in speech production.

It may be that even with absolute sameness of consonant articulation and voice quality the absolute degree and duration of pitch perturbation is variable, and hence language specific (though, on the other hand, such language specific differences may often be due to the effect of additional articulations, e.g. glottalization). To the extent that pitch perturbation is an independent variable it must compete with the exploitation of Fo for other purposes than cueing the category membership of initial stops, e.g., one may expect the duration of such perturbations in postconsonantal vowels to be shorter or less pronounced in languages with lexical tone. There is very little data, and it is not clear to what extent there is empirical support for that (cf. Hombert, Ohala and Ewan 1979, p. 41 with references). This line of reasoning, however, rather speaks against pitch perturbation as the origin of tone split (as in Thai), whereas it would be favourable to the explanation of how non-tonal languages may become tonal, cf. Modern Tibetan or the tonal dialect of Kammu (Khmu), to mention a Sino-Tibetan and an Austroasiatic case.

With these reservations in mind one must concede that numerous languages - even such that are already tonal - seem to strongly corroborate the idea that pitch perturbations caused by initials may develop into tonal contrasts (along with a loss of manner contrasts in the initials) so as to create the situation referred to by Brown (1975) as "voiced-low", viz. lower (or rising) tone after a formerly voiced initial, and higher (or falling) tone after a formerly voiceless initial, schematically: *baa > pàa as against *paa > páa.

II. THE SITUATION IN THAI

Also within the Tai family there are languages and dialects following the pattern outlined above. From a general linguistic point of view the dialects within Thai proper are particularly interesting in this context because the Siamese writing (dating back to the late 13th century) and old textbooks give very direct and interesting information about past stages. It seems rather well established (both from the spelling of Indic loanwords, from comparative studies within Tai languages, and from the classifications of the consonants in old textbooks) that there used to be a fourfold manner distinction in stops and at least a twofold distinction in other types of consonants, viz. (1) aspirated voiceless stops or continuants (called "high"), (2a) unaspirated voiceless stops, and (according to Li's reconstruction) (2b) glottalized voiced stops ((2a) and (2b) being both referred to as "mid"), and finally (3) plain voiced stops (called "low"; these are now aspirated voiceless stops, e.g. in Central Thai, having merged with category (1)). Thus, for exemplification: (1) *ph and *hm (or *m̥), (2a) *p, (2b) *ʔb, (3) *b and *m.

It is in itself a highly controversial issue what the terms "high", "mid", and "low" referred to originally, although it is off-hand tempting to relate them to a pitch-perturbating

effect (we expect fully voiced consonants to condition a lower pitch in the beginning of the following vowel than other consonants, and although it is not universally so that aspirated voiceless stops condition a higher pitch than other consonants, this may well have been the case in Old Thai). If one takes a look at the oldest textbook on Thai spelling and Thai verse (Phra Hora Athibodi's Cindamani or Cindamuni from the second half of the 17th century) one is indeed intrigued by the terminology. The letters that are still today referred to as "high" consonants are said to have a high "siang" (sound), all other consonants are said to have a mid (or medium) "siang" - this may be construed to have to do with Fo, but I presume that it rather refers to the presence versus absence of a high-pitched noise component. The consonants of mid "siang" are in turn divided into a set (the "low" consonants) that are said to have a low "kong", and the rest (the "mid" ones), which are said to have no "kong". The word "kong" is suggestive of a resonating, bell-like sound and must refer to the low-pitched full voicing which at one time was associated with "low" consonants; it is perhaps somewhat more surprising that no such voice is attributed to the "mid" stops reconstructed as voiced glottalized: *ʔb, etc. Maybe the voice really was less prominent here. - As I see it, it is not all that clear to what extent these pronunciations prescribed in a normative 17th century textbook reflect the everyday language of the area of that time, and to what extent they represent an artificial, learned recital of the individual letters of the alphabet based on the philological tradition associated with Indic loanwords in Thai. Thus the evidence can hardly be taken as proof that "low" *b etc. were still voiced stops in Thai proper at that time. Anyway, the terminology (for what it is worth without painstaking interpretation) suggests that pronunciations agreeing well with current reconstructions were at least known by learned people living some 300 years ago in the former capital of Sukhothai in Northern Thailand and taught by these to the administrators in the capital of Ayudhaya farther south. (Otherwise, the information - especially on pitch in the oldest Cindamani - is too controversial to be further dealt with here.)

As mentioned already: what we look for in tonogenesis having to do with initial consonants is evidence of a "voiced-low" development. Unfortunately, as demonstrated in the literature of the last decades, there are several Tai languages and dialects, including Modern Central Thai (Modern Siamese), that are not at all of the "voiced-low" type but rather show the opposite pattern. It cannot be stated quite briefly in what sense the pattern is opposite, since the actual tone contours depend on syllable structure and old prosodic distinctions. As for Thai, the old orthography distinguishes prosodic categories by tone marks such as "máj ẽek" and "máj thoo" (a third prosodic category being simply unmarked).

Now, to take Central Thai for illustration, the tone contour on a "live" syllable (i.e. an open syllable or a syllable ending in a resonant) without tone mark starts on a low pitch

(and rises) if the initial consonant is "high", but it starts on a higher (so-called mid) pitch after a "mid" or a "low" consonant. "Live" syllables with "máj ẽek", and likewise "dead" syllables (i.e. such that end in a stop), are low-pitched (so-called low tone) if the initial is "high" or "mid" but start on a high pitch (with or without a fall to low pitch, depending on syllable length) if the initial is "low". The relationship to initials is no more transparent in syllables with "máj thoo", since in Central Thai these all start on a high pitch but either stay high or exhibit a subsequent fall to low pitch, depending on whether the initial is "low" or belongs to one of the other categories (i.e. "high" or "mid"). One possible generalization is that in Central Thai "high" initials are accompanied by a feature of low pitch (somewhere in the syllable), whereas "low" initials are accompanied by a feature of mid or high pitch. ("Mid" initials side with "low" ones in tonally unmarked "live" syllables but with "high" ones elsewhere.) Although this pattern is pretty complex and requires a graphical representation in tone charts to make it stand out more clearly (see below), it certainly reveals a correlation of old voiced initials with a feature of high pitch, and reversely for old voiceless initials, i.e. dialects such as Central Thai are, in Brown's terminology, "voiced-high". - This I shall refer to as the tone split paradox.

Choosing Modern Central Thai rather than other dialects within Tai to illustrate the tone split paradox is well motivated. The acoustic phonetics of Central Thai has been studied so extensively, especially with regard to initial stop consonants and tones (by Abramson and others) that probably no other Southeast Asian tone language or dialect can be approached on the basis of existing phonetic knowledge of similar quality and detail. Furthermore, Modern Central Thai presents the paradox in its clearest possible form. However, in order to give some notion of the intriguing situation across dialects of Thai, I have entered the tonal reflexes of three different dialects (for some syllable types) in figure 1. The chart shows three types of syllable prosodies/terminations for Proto-Thai, viz. unmarked "live" syllable, "máj ẽek"-syllable, and "máj thoo"-syllable ("dead" syllables are left out for simplicity of exposition), and each of these is combined with four types of initials, each symbolized by a representative (for "high", "mid voiceless", "mid glottalized voiced", and "low" consonants, respectively - note that "low" b- is now ph- or p- with a voiceless consonant in all dialects, whereas nasals and liquids, such as m-, l-, are preserved and give direct testimony of the voiced status of old "low" consonants).

The dialects shown are limited to just three for which I have first-hand experience with tones and their phonetic characteristics, viz. the eastern type of Northern Thai (from the start a language of its own: Khammyang), Central Thai, and Southern Thai as spoken on the West coast in the extreme South. The tonal labels are chosen so as to give a reasonable fit both

Prosodies associated with "live" syllables in Old Siamese

Initial:		zero	1 ("máj èek")	2 ("máj thoo")
H	*ph-, *hm-, etc.	N: lo ris C: lo ris S: hi-mi ris	N: mi C: lo S: hi-mi ris	N: hi, gl C: hi fall, gl S: mi
M_1	*p- etc.	N: lo ris C: mi S: lo-mi ris*	N: mi C: lo S: lo-mi ris*	N: hi, gl C: hi fall, gl S: mi
M_2	*ʔb- etc.	N: mi ris C: mi S: lo-mi ris*	N: mi C: lo S: lo-mi ris*	N: hi, gl C: hi fall, gl S: mi
L	*b-, *m-, etc.	N: mi ris C: mi S: mi fall	N: mi fall C: hi fall, gl S: lo ris*	N: hi fall, gl C: hi peak, gl hi ris, gl S: lo

Figure 1

Some tonal reflexes in Thai dialects

N = Northern Thai as spoken in the Lampang area.
C = Central Thai as spoken in the Bangkok area.
S = Southern Thai as spoken in Satun.
"fall" = falling; "gl" = irregularity of glottal vibration; "hi" = high; "hi-mi" = higher mid; "lo" = low; "lo-mi" = lower mid; "mi" = mid; "peak" = peaked: with a (small) rise+fall; "ris" = rising.
Asterisk: analysis uncertain.

*Examples of the segmental reflexes (in Central Thai) of old initials (all examples given here having old "zero" prosody): H */phi:/ > /phi:/ 'ghost'; */hmi:/ > /mi:/ 'bear'; M_1 */pi:/ > /pi:/ 'year'; M_2 */ʔdi:/ > /di:/ 'good' (minimal pair with */ʔb/ missing here); L */bi:/ > /phi:/ 'to be fat'; */mi:/ > /mi:/ 'to have'. Note that word types such as /phi:/ 'ghost' and /phi:/ 'to be fat', or /mi:/ 'bear' and /mi:/ 'to have', contrast tonally in all three dialect areas.*

to my own impressionistic data and (especially with regard to Southern Thai, for which my own data is very meagre) to the tone charts in Brown (1965). I deviate from Brown on one point for Northern Thai: I hear (and observe on Fo-tracings) a high, strongly falling tone in "máj thoo"-syllables with initial "low" consonant, where Brown indicates a mid falling tone. (As for Southern Thai, I am not sure of the nature of the distinction Brown makes between two kinds of rather low rising tones in Satun Thai, and I simply use a somewhat arbitrary labelling: lower-mid rising vs. low rising to reflect it, but that is immaterial to the general point to be made with the use of this chart.) - The two options given for Central Thai "máj thoo"-syllables with a "low" initial reflect a difference between age groups (cf. Henderson 1982).

It will be apparent from figure 1 that it is, on the one hand, hard to set up common denominators for the tonal reflexes of "zero", "máj ēek", and "máj thoo" across dialects, and, on the other hand, equally difficult (if not more difficult) to set up common denominators for the tonal reflexes of the four types of initials across dialects. There is some suggestion in the chart that "máj ēek" has, relatively speaking, been associated with some kind of lowering compared to zero tone (this is in part contradicted by Southern Thai, but that has to do with a merger of zero- and "máj ēek"-prosodies in syllables with a non-low initial). There is definitely evidence for an association of "máj thoo" with relatively high tone followed by glottalization (but again this is contradicted by Southern Thai). - The Central Thai falling reflex of "máj ēek" (with a low initial) is clearly a coalescence with "máj thoo" (with a non-low initial).

As for the initials, there is some indication that the higher one gets on the scale from "low" to "high" consonant the lower is the tone (in some sense), but this does not hold for "máj thoo"-syllables in Northern and Central Thai, and in Southern Thai we find the opposite in "máj ēek"-syllables! There are (in part ingenious) explanations available for much of this, but off-hand the pattern does not seem particularly inviting from the point of view of deriving Modern Thai tone contours from the old syllable prosodies by any kind of simple algorithm, let alone a phonetically plausible one. Most of the chart runs counter to the expected direction of Fo perturbation. The only evidence in favour of "phonetically natural" Fo perturbation by initials is the relative lowness associated with "low" consonants in Southern Thai, but then again: why is the tone of zero-syllables falling with "low" but rising with "high" consonants?? If initials mechanically cause a perturbation of the initial part of the tonal contour we should expect the opposite. Thus, there must at least be some additional mechanism (be it articulatory or perceptual) at work.

III. EXPLANATIONS OF THE TONE SPLIT PARADOX IN THAI

There are several ways in which explanations of the tone split paradox can be attempted. The strongest hypothesis is that the speech production mechanism per se is solely responsible for the nature of the tone split via pitch perturbations stemming from the intrinsic properties of the initial consonants, but that these perturbations may - for physiologically explicable reasons - deviate from the a priori expected pattern in CV-sequences (voiced-low vs. voiceless-high) and may eventually lead to the opposite, viz. a voiced-high situation. An explanation along these lines was suggested some twenty years ago by Brown (1965), his assumption being that there will be a greater or lesser force stretching the vocal chords and the underlying thyro-arytenoid muscles depending on the behaviour of the arytenoid cartilages: the pull is strongest in consonants spoken with a wide open glottis such as aspirated stops, and weakest in fully voiced consonants. For the old "mid" initials in Thai Brown posits a laryngeal gesture which should give an intermediate degree of pull. If, now, the thyro-arytenoid muscles resist the pull by different degrees of contraction, and this contraction spills over into the following part of the syllable, the effect will be high, mid, and low pitch, respectively, with the old "high", "mid", and "low" initials in Thai. If, however, the muscles give in to the force operating on them, they must relax most for the articulation with open glottis and least for fully voiced articulation, the effect being low, mid, and high pitch, respectively.

Brown's explanation (which was apparently overlooked by Hombert, Ohala and Ewan 1979, cf. their references p. 38 bottom) is an ingenious one and interesting in the rigour and explicitness with which it makes appeal to laryngeal mechanisms in explaining prosodic sound changes. Nonetheless, it should be noted that we have no independent evidence for the contentions it makes about alternative configurations giving rise to different tones in different dialects. Moreover, these contentions are clearly interpetable only for aspirates, which do in fact seem to behave rather differently in different languages (although the underlying mechanisms are not yet well understood). As for fully voiced consonants, it is unclear to me why these would have a different degree of pull and hence different pitch-perturbating effects in - as it were - inverse relationship to the way the aspirates behave: why would the alleged mechanisms tend to place the aspirates as one extreme, and the fully voiced consonants as the other extreme along the scale of pitch perturbations?

Another category of possible explanations (to some extent already suggested for Tai) involves the assumption that the manner distinctions in initial consonants did not cause tone split directly via Fo-perturbations but that features of the consonants caused voice quality differences (cf. the so-called "register" differences in several Mon-Khmer languages) in the

vowels or in the whole remaining parts of the syllables, and that these derived features of articulation in turn conditioned pitch distinctions to arise. The resulting pitch contours might then be of quite a different kind than those derivable directly from the Fo-perturbating influence of initial consonants. This is a very interesting category of explanations; unfortunately, it is hard to provide positive evidence for the alleged intermediate steps, especially in languages which - like Thai - do not exhibit dialects or attested earlier stages with "register" conditioned by initials.

Then, thirdly, one may assume that the old consonant categories did in fact trigger the tone split by causing straightforward pitch perturbations (of the types expected) in all the dialects, so that these all had the potentials of becoming "voiced-low", but that other factors intervened and more or less changed the incipient pitch movements beyond recognition. Let us assume that the tone split always started with the phonetically expected pattern: relatively lower pitch after formerly voiced than after formerly voiceless initial, the resulting tones being at first direct reflexes of these relative pitches and only "changing place" later. It seems quite possible that, at the time when the initial Fo perturbations developed into a tone split, the tonal difference between the three prosodies marked by zero, "máj ēek" and "máj thoo" was, at least in part, a difference of level ("máj thoo" marking the highest tone). Now, as stated by Gandour (1974, p. 348), "A (...) plausible hypothesis, for the Tai language family at least, is that lexical contour tones develop from already existing level tones in order to maximize perceptual distance in the tone system." If, accordingly, we assume that the pitch perturbations first developed into further levels, these levels would multiply with the levels for the syllables with the three old prosodies, and the tonal space would end up being just over-crowded, thus necessitating the development of contours. In view of the complexity of the tonal pattern of Thai we would have to assume that such a differentiation of level tones into contours could happen in a variety of ways (being in part dialect specific). The resulting contours would undoubtedly in some cases cross each other (or cross a mid level tone), and as a consequence of this it might happen that certain syllable types in certain dialects came to exhibit a spurious association of relatively high pitch (somewhere in the course of the contour) with old "low" initials, and vice versa for old "high" initials. This, then, would qualify as the "voiced high" situation. - We do not know, however, that this is the way things happened (for one thing it is anything but certain that the tones from the old prosodies were essentially level tones). - Alternatively, as suggested by Nina Thorsen, personal communication) one might in principle imagine that the rising and falling slopes after different initials developed directly into corresponding contour tones, which later simplified into level tones in some instances. Thus, it would be logically possible for syllables with old voiced initial and "máj thoo" to first develop a strongly rising tone which later simplified into a high (peaking) tone. This line

of explanation may seem less plausible in the context of the entire Thai scenario, however.

Within this third category of explanations an alternative version says that even with initial pitch perturbations conforming to the voiced-low, voiceless-high scheme, the resulting tonal pattern may well deviate from this scheme from the very start of the perceptual reinterpretation procedure leading to phonemicization of tone ("tonogenesis"). Lea (1973, p. 64) speculates what might happen if the initial Fo slopes were translated into high or low tones: "Since Fo rises after voiced consonants, the following tone may appear high since the slope into it is upward. Similarly, the fall after an unvoiced consonant may lead to an interpretation of the following tone as low." This formulation differs significantly from the just mentioned suggestion by Nina Thorsen in that the perceptual restatement from falling to low, or from rising to high, is taken to coincide with the phonemicization rather than being a later reshaping of the tones. (Lea does not himself believe in this conjecture. He finds that "a more reasonable conjecture would be that the relative Fo values within the vowel (and not the preceding slope into the vowel) affects its interpretation as a high versus low tone", which may be construed to agree with the preponderance of low tone from voiced initial vs. high tone from voiceless initial among the cases of tonogenesis cited in the general literature. Gandour 1974, p. 348 finds the latter conjecture dubious on the basis of measurements of Fo in initials in syllables with different tones.)

Brown (1975) some ten years ago replaced his earlier physiologically based hypothesis by a hypothesis of this "third" category to account for the "voiced-high" dialects, although strictly speaking this is a hypothesis involving specific assumptions about both production and perception strategies. His first claim is that syllable initial pitch lowering after voiced initials tends to be compensated for by higher pitch in the last part of the syllable, compared to the tail of the pitch contour in syllables with voiceless initials. If, now, speakers listen to the whole pitch contour, the integrated pitch impression will be similar for both syllable types, but if the listener narrows his focus down to part of the syllable, the impression of pitch may be different depending on what part of the syllable is in focus. Thus the pitch will be perceived as either lower or higher with a voiced initial than with a voiceless initial, i.e., there is a possibility for one and the same type of pitch perturbation to lead to either the voiced-low or the voiced-high situation. Brown himself has characterized his second hypothesis as "sheer speculation", but more recently Strecker (1979) has spoken in favour of it from a general linguistic point of view (without taking a stand on its validity vis-à-vis current phonetic theory). - It should be added that both Brown and Strecker emphasize the importance of other driving forces in the tonal development, viz. such that have to do with the enhancement of contrasts within the tone system, or with alleged general tendencies for pitch contours within a tonal system to be shaped

in relation to each other (e.g. Strecker's principle that a high-falling tone falls more than a low-falling tone), and that they do this with a view to the complex developments in the various dialects for which they account so impressively.

As a phonetician one is certainly struck by the elegance and explanatory power of Brown's second suggestion, but one cannot fail to question its empirical and theoretical status vis-à-vis current general phonetic research. This applies to both of its components: production and perception. As for production, I do not know of solid evidence for the claim that the pitch rise after voiced initials tends to reach a level above that of the pitch after voiceless initials. Most studies indicate that Fo after voiced stops remains somewhat lower than Fo after voiceless stops throughout most or all of the vowel, though the Fo traces may cross in individual cases (cf. e.g. figure 2 in Hombert, Ohala and Ewan 1979, p. 40). Brown himself refers to graphs of English and Modern Standard Thai, but I think his interpretation of these is controversial (note also that Lea (1973), to take one source he quotes, does not draw the conclusions drawn by Brown on the basis of Lea's graphs). As for perception, general phonetic theory may not exclude the possibility of listeners narrowing their focus down to part of the pitch contour, but there is hardly any empirical evidence so far that this is the way dialects become tonally different, if such perceptual strategies are at all available in linguistic performance.

In short, like Brown's earlier hypothesis the more recent one attempts to account for developments in the past for which we have only indirect evidence, by reference to alleged phonetic mechanisms for which we have so far neither a comprehensive general theory, nor indisputable empirical evidence from directly accessible languages. This is obviously not an altogether desirable situation, although it should not be overlooked that if Brown's account is open to criticism it is exactly because he takes the demands for explicitness and rigour in the advancement of explanations more seriously than linguists often do. It is of paramount importance both for the progress of historical linguistics and for the progress of general phonetic theory that these disciplines be coupled together. Well attested cases of sound change provide an excellent testing ground for general phonetics, and conversely, historical linguistics should employ the most recent advances in phonetic theory (rather than some 19th century notions about phonetics) to ensure lasting progress. The tonogenesis issue is a good case in point, and seen in this general perspective Brown's use of highly sophisticated phonetic argumentation is indeed laudable, but it goes without saying that if both the data (viz. unrecorded changes taking place in the past) and the theory adduced to define and explain the data must be construed to fit each other, we are on shaky ground. I think it is important to remember Brown's very cautious remarks about the speculative nature of his own hypothesis.

In a more general sense Brown may be right in tying the explanation of the tone split paradox to the interplay between the production of perturbated pitch curves and the perception and categorization of these. To illustrate the possibility of advancing alternative hypotheses within this realm of thinking, I shall permit myself to refer to a suggestion which was ventured by John Ohala in discussions with me about the tone split paradox. His suggestion (which certainly was not meant as a serious theory about tonogenesis in Thai but just as a technically possible type of argumentation) hinges on the tone level contrasts already existing in the language. Assume that there is a certain frequency value which is the boundary (threshold) between what is categorized by the listener as a lower and a higher phonemic tone. A listener hearing a syllable whose pitch is only just below this boundary frequency may categorize it as having the lower tone, or he may construe it to be a higher-tone syllable if the initial consonant is of a type (voiced stop, for example) which might be responsible for a perturbation lowering the pitch across the boundary. Assuming that the syllable was actually intended as having the lower tone, we would have in the latter case a faulty categorization on the part of the listener which might eventually lead to tonal restatement for the lexical item in question. Something similar (mutatis mutandis) might conceivably happen with syllables spoken on pitches just above the boundary value. - Obviously, this kind of reasoning can only explain how some words of the language may switch tone due to a perceptual strategy at variance with the intentions of the speaker; it remains to be explained how such a tonal switch might be generalized to all words with all types of syllable structure and all the tonal categories of Old Siamese. The ultimate interest of this specific line of argumentation depends on how safely we can predict the occurrence of tonal recategorizations near the boundary values between contrastive tones; there is some slight evidence from data in a recent paper by Abramson and Erickson (1978) which may be construed to suggest that occasional perceptual switches in categorization having to do with the nature of the initial consonants do indeed happen near tonal boundaries in Modern Central Thai. However, as emphasized by Ohala himself, it is not the case that there is solid support for the specific hypothesis outlined above, which - as said already - was advanced only as an illustration of a certain category of putative explanations.

IV. THE TONE SPLIT PARADOX IN A WIDER PERSPECTIVE

As I see it, the essential thing about such proposals is that they make specific claims about phonetically and structurally defined situations in which a given type of sound change (e.g. voiced initial giving high pitch) may take place, i.e., they are predictive in this restricted sense (not, of course, in the sense that they predict whether the sound change will actually take place in a given language or dialect at a given time).

This means that they are open to empirical testing of a weak kind: if in the languages of the world there are many languages exhibiting the change under the specific conditions set by the hypothesis but few or no languages exhibiting it in which these conditions are not met, then obviously the typological findings corroborate (though not prove) the hypothesis, and vice versa. Now, if I understand both Brown's and Ohala's arguments correctly, it seems that they make different predictions: the latter definitely sets off tone split from other types of tonogenesis and accounts for the voiced-high situation only as a special case of tone split; Brown, however (both in his first and in his second hypothesis) posits phonetic mechanisms and strategies which in principle should be able to create the voiced-high situation irrespective of the previous status of the language as either tonal or non-tonal.

Hombert, Ohala and Ewan (1979, p. 54) speak of "the occasional correlation between higher tone and originally preceding voiced consonant (vs. lower tone and voiceless consonant), where just the reverse would be expected". The area notorious for such exceptions is Southeast Asia, and the languages or dialects in question typically belong to those language families which have had prosodies (tones or "terminals") long before the initials affected the tonal pattern, i.e., these are cases of tone split. Let us imagine first that a careful search for divergent tone patterns in remote dialects of newly tonal languages (e.g. of the Tibeto-Burmese and Mon-Khmer families) would reveal that a sizeable proportion of these dialects are in fact voiced-high just as is the case in Tai. That finding would seriously detract from the explanatory power of hypotheses which - like Ohala's suggestion - are crucially sensitive to the initial tonal conditions of the language undergoing tonogenesis. If, on the other hand, it is true that newly tonal languages in which tones reflect the previous voicing state of initials are (with few and marginal exceptions) voiced-low across the board, this spells trouble for the category of explanations represented by Brown's first and second hypotheses. It is not that these are shown to be wrong, but it takes a powerful, additional explanatory device to account for the typological difference observed in a wider array of languages.

Thus, quite apart from considerations of the inherent phonetic plausibility of one or another hypothesis, we have a statistico-typological approach to the general issue. As for the tone split paradox, one thing that might be done to get a more solid empirical foundation for the formation of general phonetic hypotheses about tonogenesis is to search carefully for tone languages or dialects all over the world which seem to have developed high-pitched tone from voiced initials, and to see whether these are practically all instances of tone split (except for cases which obviously have special explanations), or whether there are also clear-cut instances of non-tonal languages becoming tonal via voiced-high tonogenesis.

Unfortunately, the situation is not as simple as suggested above with regard to the possibility of revealing universal tendencies in tonogenesis under specified structural conditions. If we find a statistical preponderance of a certain phenomenon, or of a correlation among certain phenomena, that in itself does not tell us anything about universals of human languages, of course. We have to know to what extent the phenomenon or correlation in question is a truly universal tendency and as such triggers mutually independent spontaneous developments, and to what extent the various occurrences we have observed, are linked by some kind of interrelatedness among the languages or dialects under study. This is a rather trivial statement, to be true, but it may be appropriate to consider the problem in its general outline here since its importance is not always made explicit in connection with purely theoretical argumentation about the mechanism(s) of tonogenesis.

Languages may be interrelated by belonging to the same family or sub-family, i.e., by genetic relatedness, which in the strict sense implies that the correspondences among them are consistent with the assumption of a common parent language, though the scenario is often not all that transparent. As for South and Southeast mainland Asia, we know from authoritative comparative research that there are five well-defined families or sub-families which include tonal languages, viz. Chinese, Tibeto-Burmese, Tai (-Kadai), Mon-Khmer (Austroasiatic), and Indo-Aryan. There is no proof of genetic relatedness (so far) between families, except for Chinese and Tibeto-Burmese. In any case, genetic relatedness across the five groups enumerated here belongs to a very distant past and can hardly be directly relevant to recent tonogenesis taking place essentially within the last millennium. If, however, we look within each of the five families it may be rightfully claimed that tonogenesis in each particular language or dialect should first be viewed in a comparative perspective: as possibly derived from a common ancestral source together with tonogenesis in related languages or dialects. It may be that tonogenesis happened already in a parent language, or it may be that languages of a genetically defined group carry a structural and/or phonetic predisposition for developing in a specific direction, i.e., that we have an instance of "drift", whatever that exactly means. (Candidates for "drift" are, e.g., the tendency toward developing a binary tonal distinction caused by initials via a "register" difference, as in some Mon-Khmer languages, and the tendency toward developing tonal distinctions caused by "finals".)

This suggests, not surprisingly, that evidence for universals should include samples of languages belonging to several different families. It is equally essential, however, that there are samples from one and the same family if the languages in question have developed in significantly different ways so that they do not just appear as projections of an old, shared development or as manifestations of a common "drift" (cf. that Mon-Khmer comprises tonal and non-tonal dialects even within

one and the same language, and that this language family comprises tonal languages of the simplest possible type with a binary distinction as well as a language of a highly complex type with many contrasting contour tones: Annamese).

In addition to genetic relatedness we have the question of areal features. The area stretching from the southeastern coast of mainland Asia to the regions just north and south of the Himalayas, is today one vast tone language area. Why is there a tonal area of such extension, and why does it include several language families? In Southeast Asia there is apparently one layer of tonogenesis having to do with old finals, and another layer having to do with old initials. Why is the latter, rather than the former, manifested in recent cases of tonogenesis, and why is it more widespread geographically? Tonogenesis conditioned by initials is known outside this area, but to what extent is Southeast Asian tonogenesis from finals matched elsewhere? Is it somehow related in type to what happened in certain Scandinavian languages or dialects?

With such an areal feature as tonogenesis there may well be "drift" within each language family or sub-family, but this cannot possibly account for unrelated languages sharing such a development. How is the tendency to develop tones diffused across dialect or language boundaries? It is too far-fetched to assume that it is a coincidence that several unrelated languages of this area favour tonogenesis (cf. that Chinese, Annamese, and Tai languages share certain developments causing a high degree of tonal complexity which should be viewed in the light of the intimate and long-lasting contact known to have existed between these languages).

The general theory of language change thus faces a twofold task: to reveal and explain universal tendencies and to account for the ways in which innovations may be imported from one language into another. Phonetically (i.e. physiologically and/or perceptually) "natural" tonogenesis which occurs spontaneously and independently in many places is one thing; borrowing of tonal contours or perhaps even tonal distinctions is something very different. We do not know what it takes for a tonal pattern, or for a set of (possibly arbitrary) substitutions between segmental and tonal distinctions, to be copied from one language into another, and to what extent this is possible. "Crazy" types of tonogenesis might well have to do with language contact, e.g. borrowing of a tone contour which is introduced in certain items of the vocabulary (according to etymology) without any internal phonetic motivation and eventually changes the overall pattern.

V. CONCLUSION: THE COMPLEXITY OF TONAL DEVELOPMENTS

With our present knowledge it does not seem very likely that the tone split paradox can be explained solely on the basis of production and perception mechanisms and strategies operating on initial consonants and on pitches that are perturbated due to the presence of these consonants. The entire phonological system (the conditions to which system-internal forces respond) must be taken into due consideration when explaining the development of tones. Brown, Egerod, and Strecker attempt to do this (in various ways), but there are too many unknown factors involved e.g. with reference to the exact articulation of initials. Note, for example, that initial *b may have had a very different effect on Fo depending on whether it was at some chronological point just voiced or voiced aspirated ("breathy") or possibly implosive (Hombert, Ohala, and Ewan 1979 cite data according to which the onset Fo of the vowel after a breathy voiced consonant is markedly lower than after a plain voiced one, and other data according to which Fo after an implosive may be higher than after a plain voiced stop). - Gandour (1974) has actually suggested a reconstruction */bɦ/ at some chronological point, on the basis of evidence from Southern Thai (Egerod 1961), similarly Egerod (1971).

Moreover, there may be irretrievable differences in the relative chronology of events which have made otherwise closely related dialects develop rather differently on this point. In this context it is very important to take into consideration the role played by linguistic variation, in particular by coexisting pronunciation norms associated with different generations, different social groups, or different styles of speech. Unfortunately, there is but little available knowledge about such parameters of variation in the earlier stages of Thai, although the linguistic diversity (involving several dialects of Thai proper) is known for sure to be of considerable age in Thailand, so that bilingualism (in a wider sense) must also be considered as an essential factor in tonogenesis (possibly promoting seemingly ad-hoc system-internal readjustments such as tonal flip-flop?).

As for the dialects of Thai, these are mutually so closely related that it is rather tempting to account for their differences in tonal patterns as relatively late - though in part radical - deviations from a more or less common basis which may have included the phonetic prerequisites for development of a voiced-low tone split. But even under such an extreme hypothesis about linguistic homogeneity there are many degrees of freedom in the pattern one may posit for the old language. Take the tone marks of Old Siamese, for example: did these represent tone levels, rising or falling tone contours, phonation types involving features other than pitch, or what? The modern reflexes (cf. figure 1 above) very strongly suggest that the prosody indicated in Thai by "mâj thoo" had a component of syllable-final glottalization (or laryngealization)

all along. As for the prosody indicated by "máj ẽek", it may well at one point have had a component of final aspiration or breathiness (this is consistent with the tendency toward final low Fo in many modern reflexes of this syllable type), although Brown 1965, p. 38 & 52 just posits voicing finally in Ancient Thai. But it is hard to get beyond that.

As for the consonant categories called "high", "mid", and "low" the strong assumption of pitch perturbations corresponding to these terms does not tell us how these pitch perturbations eventually modified the existing tones (before the emergence of phonemic tone splits): did they result in pairs of allotones differing in overall level, for each tone, or was the result a split into allotones differing in the direction of the pitch movement? Did allotones differing first in pitch level later develop into contours because of influence from the rest of the tonal system? Even given a fully specified set of initial conditions we cannot fully predict what course the tonal developments will take, and of course we cannot predict whether tonogenesis will occur at all. General phonetic theory helps very much to understand what has happened, but we have to know more about the phonetics of past stages.

Basically, Marvin Brown and the other pioneers have been doing what must be done, viz. to combine scrutiny of the philological information available (from the spelling system, from old grammars, from ancient loanwords whose literary route is known, etc.) with the application of modern linguistic and phonetic theories. This line of research may be pursued further, however, especially with regard to an inclusion of the most recent work on speech physiology and perception and a reinterpretation of the evidence from old sources such as the Cindamani referred to earlier in this paper.

VI. FINAL REMARKS

Two further observations should be made concerning the relationship between initial Fo perturbation and tone: (1) It should not be overlooked that the "low" consonants of Thai include sonorants as well as (old voiced) stops. Sonorants are not supposed to lower Fo significantly but still those have exactly the same tonal reflexes as "low" stops. This suggests that the tone split should be explained with reference to the perturbating effect of non-low rather than low consonants.

(2) As mentioned earlier, there is some evidence that the duration of the Fo perturbations caused by the presence of preceding consonants is limited in languages in which the vowels carry lexical tones. Since there may nevertheless be considerable differences in vowel-initial Fo as conditioned by the preceding consonant, it seems as if the very start of the vowel is less crucial for the perception of lexical tone than the remainder. Now, if we look at the tonal patterns of modern Thai dialects (i.e. the patterns emerging from the combined effect of old final prosodies and old manner differ-

ences in initials), we may well venture the contention that on the whole it is the final part of the contours that is most important for the characterization of lexical tones in terms of meaningful features such as "rise", "fall (±glottalization)" etc. (I think this is consistent with the extensive research on the acoustics and the perception of Thai tones by Abramson, Gandour, and others). It would be interesting to know to what extent this is true generally. There is good evidence (Brown 1965, Egerod 1971) that tone in languages such as Thai is a property associated with the final part of the syllable and hence forming a joint system with such properties as checked vs. unchecked syllable, breathy vs. plain vs. glottalized vowel offglide, and long vs. short vowel. If there is in fact a clear-cut perceptual division between syllable initial properties being interpreted as non-prosodic irrespective of their phonetic nature, and syllable central and final properties being in part interpreted as prosodic, we may see the "phonologization" of initial Fo perturbations as a categorial shift from perceiving these Fo movements as part of (or coarticulation with) the syllable initial to perceiving them as part of the syllable remainder. (This formulation deliberately leaves it entirely open whether there is possibly a related categorial difference between properties associated with the vowel per se and properties associated with its termination, duration being interpretable as a "central" vowel feature in some languages and as a "terminal" feature in others, which might account for its varying status as a prosodic or non-prosodic property.)

REFERENCES

Abramson, A.S. and Erickson, D.M. 1978: "Diachronic tone splits and voicing shifts in Thai: some perceptual data", *Status Rep. Speech Res., Haskins Labs. 53.2*, p. 85-96

Brown, J. Marvin 1965: *From Ancient Thai to modern dialects* (Soc. Sc. Found. Press of Thailand, Bangkok; new edition Bangkok 1985)

Brown, J. Marvin 1975: "The great tone split: did it work in two opposite ways?", in: *Jimmy G. Harris and James R. Chamberlain (eds.): Studies in Tai linguistics in honor of William J. Gedney* (Office of State Universities, Bangkok), p. 33-48

Egerod, S. 1971: "Phonation types in Chinese and South East Asian languages", *Acta Linguistica Hafniensia 13*, p. 159-171

Egerod, S. 1961: "Studies in Thai dialectology", *Acta Orientalia 26*, p. 43-91

Erickson, D.M. 1975: "Historical implications for an historical account of tonogenesis in Thai", in: *Jimmy G. Harris and James R. Chamberlain (eds.): Studies in Tai linguistics in honor of William J. Gedney* (Office of State Universities, Bangkok), p. 100-111

Gandour, J. 1974: "Consonant types and tone in Siamese", *J. Phonetics 2*, p. 337-350

Halle, M. and Stevens, K.N. 1971: "A note on laryngeal features", *MIT Quart. Prog. Rep. 101*, p. 198-213

Hombert, J.-M., Ohala, J.J., and Ewan, W.G. 1979: "Phonetic explanations for the development of tone", *Language 55*, p. 37-58

Kohler, K.J. 1985: "Fo in the perception of lenis and fortis plosives", *J. Acoust. Soc. Am. 78*, p. 21-32

Lea, Wayne A. 1973: "Segmental and suprasegmental influences on fundamental frequency contours", in: *Larry Hyman (ed.): Consonant types and tone* (Univ. of Southern California), p. 15-70

Li, Fang-Kuei 1943: "The hypothesis of a pre-glottalized series of consonants in primitive Tai", *Bull. of the Inst. of History and Philology, Academia Sinica 11*, p. 177-188

Matisoff, J. 1973: "Tonogenesis in Southeast Asia", in: *Larry Hyman (ed.): Consonant types and tone* (Univ. of Southern California), p. 71-95

Petersen, N. Reinholt 1983: "The effect of consonant type on fundamental frequency and larynx height in Danish", *Ann. Rep. Inst. Phon., Univ. Copenhagen 17*, p. 55-86

Riordan, C.J. 1980: "Larynx height during English stop consonants", *J. Phonetics 8*, p. 353-360

Rischel, J. 1985: "An appraisal of research in the phonetics and phonology of Thai", *Ann. Rep. Inst. Phon., Univ. Copenhagen 19*, p. 43-93

Strecker, D. 1979: "Higher falls more: A tonal sound change in Thai", *Computational Analyses of Asian and African Languages 11*, p. 30-84.