

Svenska studenters uppfattningar av tonerna i kinesiska tvåstaviga ord

Guohua Hu

Centrum för tvåspråkighetsforskning
Masteruppsats, 30 hp
Masterprogram i språkvetenskap med
inriktning mot tvåspråkighetsforskning
Våren 2015
Handledare: Niclas Abrahamsson



Stockholms
universitet

Tackord

致谢 <zhixiè>

师傅领进门，修行在个人。

<shīfú lǐngjìn mén, xiūxíng zài gè rén>

Mästaren visade vägen men varje steg var min uppgift.

Svenska studenter uppfattningar av tonerna i kinesiska tvåstaviga ord

Guohua Hu

Sammanfattning

Studenter med ett icke-tonspråk som modersmål brukar ha svårigheter att lära sig de kinesiska tonerna när de börjar sina studier. Å ena sidan brukar kursböckerna på denna nivå använda bara enstaviga ord och det har ansetts vara god pedagogik att läraren visar tonkonturena med sina händer. Å andra sidan har forskningen om kinesiska länge varit koncentrerad på vilka värden som grundtonen (F0) uppvisar i enstaviga ord. Det är många faktorer som tvärspråksforskningen inte har uppmärksammat, bl.a. det inflytande som konsonanterna har på F0 (vilket infödda talare inte alltid är medvetna om).

Det finns ingen samsyn när det gäller att förklara tonförväxlingsmönstren. Tidigare teorier inom ämnet andraspråksinläring (Second Language Acquisition, SLA), som Perception Assimilation Model (PAM) och Speech Learning Model (SLM) har visat sig otillräckliga för studiet av tonperception. På senare tid har PAM-Suprasegment försökt förklara hur den lärandes modersmål antas närma sig kinesiskans tonsystem men modellen tar inte upp ordprosodin. Eftersom den moderna kinesiskans ordförråd till majoriteten består av tvåstaviga ord bör forskningen lämna gamla spår för att finna andra kriterier – som duration och betoning – för att förklara hur man lär sig höra kinesiskans toner, t.ex. vad som händer när två toner återfinns i ett ord.

Denna studie har som mål att ta reda på hur svenskar som studerar kinesiska som andra språk/främmande språk uppfattar tonerna i kinesiska tvåstaviga ord. Experimentet bygger inte på tillrättalagda testord. Resultaten visar att tonerna först och främst påverkas av den initiala konsonanten och sedan av de omgivande tonerna. Vidare visas att det svenska systemet med accent I och II i sin tur kan åstadkomma tonförväxlingar eftersom kinesiska tvåstaviga ord delvis uppvisar samma mönster.

Resultaten illustrerar att tonidentifiering är en dynamisk och komplex process. Det krävs fortsatt forskning om tonerna för att få grepp om dem men det kan inte stanna där: interaktionen mellan ljud och ordprosodi behöver belysas bättre för att uppnå god behärskning av prosodin i språkundervisningen.

Nyckelord: kinesisk ordprosodi, kinesiska toner, kinesiska konsonantkategorier, tonmissuppfattningar, kinesiska tvåstaviga ord, svenska konsonantkategorier, svenska ord-accenter

Swedish students' perception of the tones in Chinese disyllabic words

Guohua Hu

Abstract

Foreign adult students with atonal language usually have, in the beginning of their Chinese study, difficulties to identify the Chinese tones. On one side, only monosyllabic tones are mentioned in course books during this earlier stage and to illustrate the tone contours with hands has been treated as an effective pedagogy. On the other side, research on Chinese has for long been solely concentrated upon the values of the fundamental frequency (F0) of the vowels in monosyllabic words. In cross-linguistic studies many factors, among others the effects of consonants on F0 that native speakers are not aware of, have still not been paid special attention to.

There is no consensus regarding the explanation to tone confusion patterns. Earlier theories of Second Language Acquisition (SLA) like Perception Assimilation Model (PAM) and Speech Learning Model (SLM) are no longer suited for tone perception. Recently, PAM-Suprasegment has tried to approach that the intonation of the learners' native language is assumed to be assimilated to the Chinese tone system. However, this model ignores the word prosody. Nowadays, when the modern Chinese vocabulary consists of a majority of disyllabic words, research has to be re-directed to find other criteria e.g. temporal and stress for explaining the complexity of Chinese tone perception, i.e. how two tones behave when they are combined in one word.

The purpose of this essay is to explore how native Swedish speakers learning Chinese as second/foreign language perceive the Chinese tones of disyllabic words. The experiment is not based on elaborated test words. The results show that tones are first of all affected by the initial consonants and sequentially influenced by the surrounding tones with accordance to Chinese. It further reveals that Swedish accent I/II patterns might be a reasonable explanation for the Chinese tone confusion patterns since partially acoustic properties of Chinese disyllabic words overlap the Swedish accents.

These results mean that tone perception is a dynamic and complex process. Further research on tone perception should explore profoundly and widen interaction between sounds and word prosody, which paves the way for more effective prosodic practice in language education.

Keywords: Chinese word prosody, Chinese tones, Chinese consonant categories, tone confusions, Chinese disyllabic words, Swedish consonant categories, Swedish word accents

Innehållsförteckning

Tackord	2
Sammanfattning	3
Abstract	4
Tabeller	6
Figurer	7
1 Inledning	8
2 Bakgrund	9
2.1 L1 kinesiska toners akustik	9
2.1.1 Tonemens akustik	9
2.1.2 Tonsandhi akustik	11
2.2 L1 kinesiska tonperception	13
2.2.1 Tonemens perception	13
2.2.2 Tonsandhi perception	15
2.3 L2 kinesisk tonperception	15
2.3.1 Tonem	15
2.3.2 Tonsandhi perception	16
2.4 Jämförelse mellan kinesiska och svenska	18
2.5 SLA-teori och kinesisk tonperception	20
2.6 Sammanfattning	22
3 Frågeställning	23
4 Metod	23
4.1 Informanter	23
4.2 Stimuli	24
4.3 Procedur	25
4.4 Databehandling och dataanalys	25
5 Resultat	26
5.1 Tonidentifikation	26
5.1.1 Tonidentifikation i $\sigma 1$	26
5.1.2 Tonidentifikation i $\sigma 2$	26
5.2 Tonförväxlingar	27
5.3 Förhållandet mellan konsonantkategorier och tonförväxlingar	27
5.3.1 Allmänt i $\sigma 1$	27
5.3.2 Allmänt i $\sigma 2$	28
5.4 Tonkombinationer	28
6 Slutsats och Allmän diskussion	29
6.1 Slutsats	29
6.2 Allmän diskussion	29
7 Slutord och framtida forskning	32
Referenser	34
Bilaga 1	37
Bilaga 2	38
Bilaga 3	39
Bilaga 4	40
Bilaga 5	41
Bilaga 6	42
Bilaga 7	43

Tabeller

<i>Tabell 1 Samtliga möjliga tonkombinationer</i>	<i>11</i>
<i>Tabell 2 Uppgifter om informanterna.....</i>	<i>23</i>
<i>Tabell 3 Fördelningen av Stimuli (T) × Felaktiga Responser (T') i $\sigma 1$</i>	<i>26</i>
<i>Tabell 4 Fördelningen av Stimuli (T) × Felaktiga Responser (T') i $\sigma 2$</i>	<i>26</i>
<i>Tabell 5 Sammanställning av $\sigma 1$ och $\sigma 2$, Independent Sample T-test</i>	<i>27</i>
<i>Tabell 6 Tukey HSD^{a,b} av FTKR</i>	<i>29</i>

Figurer

Figur 1 Normaliserad tid (horisontellt) och medelvärden F0 (lodrätt) av stavelsen <ma> med fyra toner (hämtad ur Xu 1997:67)	10
Figur 2 F0-kurvor för σ1 (vänster: 前字四声 <qianzisishēng> 'fyra toner för det första tecknet') och σ2 (höger: 清声母后字的四声 <qīngshēngmùhòuzide sishēng> 'fyra toner med ett tonlöst ljud för det andra tecknet'; T1: ---, T2: ----, T3: -----, T4: -----) i ledfrasen 这两个字组, 您怎么说? <zhe liǎngge zìzǔ zěnme shuō?> 'Hur uttalar ni det här tvåstaviga ordet X?', en manlig pekingbotalare (hämtad ur Wu & Lin 1989:163).....	12
Figur 3 F0-kurvor för σ1 (vänster: 前字为阴平和阳平 <qianzi wéi yīnpíng hé yángpíng> 'det första tecknet med T1 och T2') och σ2 (höger: 非清声母后字的四声 <fēiqīngshēngmù hòuzide sishēng> 'fyra toner med ett tonande ljud för det andra tecknet', T1: ---, T2: ----, T3: -----, T4: -----) i ledfrasen 这两个字组, 您怎么说? <zhe liǎngge zìzǔ zěnme shuō?> 'Hur uttalar ni det här tvåstaviga ordet X?', en manlig pekingbotalare (hämtad ur Wu & Lin 1989:163)	12
Figur 4 F0-kurvor för σ1 (vänster: 前字为上声和去声 <qianzi wéi shǎngshēng hé qùshēng> 'det första tecknet med T1 och T2') och σ2 (höger: 非清声母后字的四声 <fēiqīngshēngmù hòuzide sishēng> 'fyra toner med ett tonande ljud för det andra tecknet', T1: ---, T2: ----, T3: -----, T4: -----) i ledfrasen 这两个字组, 您怎么说? <zhe liǎngge zìzǔ zěnme shuō?> 'Hur uttalar ni det här tvåstaviga ordet X?', en manlig pekingbotalare (hämtad ur Wu & Lin 1989:163)	13
Figur 5 Timing of turning points för <ai> (vänster: T2 före 400 ms respektive höger: T3 efter 400 ms, hämtat ur Shen & Lin 1991:147)	15
Figur 6 Jämförelse mellan svenska och kinesiska klusilerna enligt Catford's modell (hämtad ur Hu 2012).....	18
Figur 7 Vågformer och F0-kurvor för orden Milan (entoppig) och milan (tvåtoppig) i ledfrasen Det var_. Manlig talare (hämtad ur Engstrand 2004:187).....	20
Figur 8 Vågform (överst), spektrogram (mellan) och blåmarkerat F0-kontur av ett test ord nr. 1-黄河- <huánghé>-[xwanʅ-xʂʅ] 'Gula Floden'. Manlig respektive kvinnlig talare.....	24
Figur 9 Förhållandet mellan stimuli och konsonantkategorier i σ1	27
Figur 10 F0-kontur (överst), IPA (mellan) av testorden nr. 13-新奇-<xīnqí>-[ɕinʅ-tɕʰiʅ]- 'nyfiken'; manlig talare respektive kvinnlig talare och nr. 17-终于-<zhōngyú>-[ʈʃunʅ-yʅ]- 'äntligen'; manlig talare (till vänster) respektive kvinnlig talare (till höger).	31
Figur 11 F0-kontur (blå heldragen linje) av testorden nr. 23-托运-<tuōyùn>-[tʰuoʅ-ynʅ]- 'frakt'; manlig talare respektive kvinnlig talare och nr. 7-开阔-<kāikuò>-[kʰaiʅ-kʰwoʅ]- 'vidsträckt'; manlig talare respektive kvinnlig talare	31

1 Inledning

Vuxna som har ett icke-tonspråk som modersmål har svårt att särskilja de kinesiska tonerna. Forskningen på området har haft olika utgångspunkter, bl.a. de kinesiska tonernas akustiska egenskaper, tonperception och psykolingvistiska perspektiv på tvärspråksforskningen.

Betydelsen för kinesiska enstaviga ord ändras genom att (T)oner (vokalens grundton, F0, *Fundamental frequency*) ändras. Det är F0 som avgör vilken ton stavelsen uppfattas ha (Yip 2002:292). Som ett resultat av detta föreligger problemet att tvärspråksforskningen mest sysslat med syntetiska enstaviga ord där F0 har extraherats. Från extremtester finns det även tecken på att amplituden (Whalen & Xu 1992) och tonkurvornas inriktning (Gandour 1984) i viss mån spelar in för personer som har kinesiska som modersmål. Det saknas däremot nästan helt tvärspråksforskning angående hur perceptionen av F0 påverkas av andra faktorer, som t.ex. konsonanter.

För vuxna studerande som inte är kinesisktalande har forskarna (t.ex. Kiriloff 1969) funnit att av de fyra kinesiska tonerna är T2 svårast att uppfatta medan T4 är lättast och att tonförväxlingar oftast sker mellan T2 och T3. Särskilt lätt att uppfatta är den fallande T4 när den uppträder finalt i ett yttrande på samma sätt som i engelskan (Broselow et al. 1987). Svårigheterna att identifiera T2 och T3 beror på att deras akustiska egenskaper liknar varandra (Chuang et al. 1972). En annan fonetisk förklaring är att *timing of turning point* (Shen & Lin 1991) skulle kunna vara den avgörande faktorn för att skilja T2 från T3; detta testades dock enbart med tonande initiala konsonanter så frågan är om resultaten står sig när/om det är tonlösa initiala konsonanter som inleder testorden. Härav kan dras slutsatsen att valet av konsonant/vokal för experimenten är viktigt, en faktor som inte har uppmärksamats speciellt.

Förutom tonakustik visar Hallé et al. (2004) ur en psykolingvistisk vinkel att icke-tonspråkstalare saknar de fonologiska strategierna för att kategorisera kinesiska toner. So & Best (2014) bekräftar, genom Perception Assimilation Model-Suprasegment (PAM-S) baserad på PAM (segment, Best 1995; jfr. Flege 1995) att engelsktalande i T4 tycker sig höra den intonation som föreligger i engelska påståendesatser (se ovan) medan fransktalande upplever sig känna igen den intonation som förekommer i utrop på franska. De menar att modersmålets intonation assimileras till det kinesiska tonsystemet.

Ordprosodi är inte detsamma som intonation men detta har man lämnat oproblematiserat när man låtit toner komma i kontakt med varandra (t.ex. Lee et al. 2010). Kinesisk lexikologi består till en kraftig majoritet av tvåstaviga ord (Wang et al. 2004:194). Deras ordprosodi skiljer sig från de enstaviga ordens (Duanmu 2007:129-158) på så sätt att även övriga prosodem (temporala och dynamiska) får viktigare roller (McAllister 1998:134-136). Tonuppfattning (även gällande enstaviga) är alltså komplex, dynamisk alltifrån segments inverkan till att en ton påverkar en annan ton över stavelsegränserna.

Denna studies mål är att komma fram till om tidigare utländska resultat också gäller svensktalande och att i så fall undersöka vad som förorsakar att svenska vuxna lyssnare missuppfattar kinesiskans toner. Det finns bevis för att tonande respektive tonlösa konsonanter i engelska, franska och yuruba inverkar störande på F0 (Hombert 1978) – misstanken finns att detsamma kan föreligga i mötet mellan svenska och kinesiska. De kinesiska aspirerade konsonanterna stör F0 akustiskt (Xu & Xu 2003). Uppsättningen av svenska tonlösa klusiler skiljer sig från vad som föreligger i kinesiskan (Hu 2012) så det är berättigat att fråga sig hur mycket kinesiskan inverkar och hur svenskans tonaccenter (Bruce 2005; Engstrand 2004:186-197; Gårding 1998) överförs till kinesiska.

Materialet tillhandahölls av en lärare i kinesiska med lång erfarenhet och var inte tillrättat för analyser av det slag som presenteras i detta arbete. För att undersöka den dynamiska processen samt få en mer objektiv bild genomfördes samma test med deltagarna från nybörjarkursen i kinesiska tre år i följd. Testorden är tvåstaviga. Studien är avsedd att vidga vår kunskap om och förståelse för hur studerande med explicit kunskap kan tänkas tillägna sig tonerna i ett L2; härigenom skulle litteraturen om både akustisk teori och PAM(S) utökas.

Fynden kan ge förslag till pedagogisk verksamhet där lärarnas gestikulerande fortfarande dominerar när de beskriver de kinesiska tonerna (Chen 2013b).

Dispositionen är som följer: §2 behandlar tidigare studier om tonperception av både en- och tvåstaviga ord för både L1- och L2-talare; i §3 presenteras en hypotes baserad på tidigare forskning; §4 redogör för den aktuella undersökningen och §5 för dess resultat; §6 innehåller slutsatser och allmän diskussion medan slutligen §7 antyder vägar för pedagogik och vidare forskning.

2 Bakgrund

Varje kinesisk stavelse (motsvarande ett kinesiskt tecken) består i princip av tre lika viktiga element: konsonant, vokal och ton (Wang et al. 2004:81). Stavelsen <ma> avviker betydelsemässigt från <mi> och <na> dels genom initialfonemets uttal [m/n], dels genom de skilda vokaler [a/i] och dels får <ma> olika betydelser beroende av vilken ton (T) stavelsen bär: 妈 <mā> 'mamma', 麻 <má> 'hampa', 马 <mǎ> 'häst' och 骂 <mà> 'skälla' samt den neutrala tonen som återfinns i den tonlösa frågemarkören 吗 <ma>.

Ton(em), "[B]y analogy with phoneme [...] can serve alone to distinguish words" (Trask 1995:357), d.v.s. utgöra minsta betydelseskiljande enhet (Sjögren 1979:147). Hallé et al. (2004:396) hävdar också att "[T]onal information is perceived by native speakers of tonal languages in a linguistically contrastive way, presumably just like segments are". Av detta framgår att det är av yttersta vikt att kunna både uppfatta och återge tonerna på rätt sätt. För en fungerande kommunikation räcker det dock inte att bemästra tonen i enskilda stavelser – detta liknar resultatet när kineser, som läser på ett främmande språk, enbart skiljer fonem från fonem helt utan ordprosodi och intonation. Dessutom är mer än 70 % av de moderna orden i kinesiska minst tvåstaviga (och omfattar alltså minst två tecken, Duanmu 2007:160; Wang et al. 2004:194). I kontext, d.v.s. när ton kombineras med annan ton, förändras de. Då gäller det inte enbart F0 utan även andra prosodem t.ex. temporala och dynamiska drag (McAllister 1998:134-136). För att reda ut vad som krävs för att uppfatta tonerna behöver vi förstå deras akustiska egenskaper.

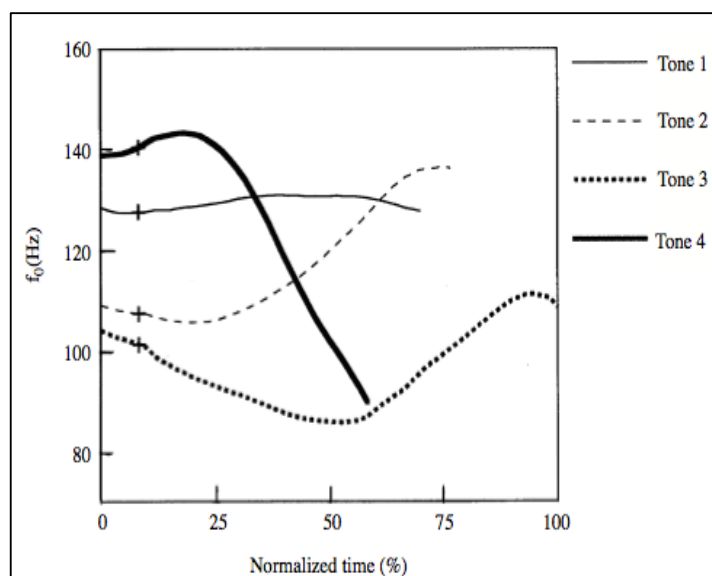
Mot denna bakgrund blir det av vikt att presentera tonemen och tonkombinationerna var för sig med avseende på både tonernas akustik och perception ur både modersmåls(L1)- respektive målspråkstalandes (L2) perspektiv.

2.1 L1 kinesiska toners akustik

2.1.1 Tonemens akustik

Grundtonsfrekvens, (F0, *Fundamental frequency*), som mäts med Hz, är den frekvens med vilken stämläpparna vibrerar per sekund. Upplevd tonhöjd (*pitch*) använder mel-skalan som är logaritmisk; dess relation är inte linjär till F0, men det är ändå frekvensen som bestämmer tonhöjden (Engstrand 2004:260; Guo 1993:73; Trask 1995:278; Yip 2002:289). Tonkurvorna är inte så stabila som de diakritiska markeringarna antyder utan kan variera under förloppet (durationen, uttrycks i millisekunder, ms), jfr. figuren nedan. Med amplitud avses röststyrkan som mäts med decibelskalan (dB).

Akustiska analysinstrument (t.ex. Praat www.praat.org) hjälper oss illustrera och utforska dessa tonegenskaper. Figur 1 visar relationerna mellan ton, tonhöjd, duration och tonkurvor. Lodrätt redovisas i figuren F0 (Hz) och horisontellt durationen för de fyra tonerna (T0 är inte medtagen) i det enstaviga <ma>.



Figur 1 Normaliserad tid (horisontellt) och medelvärden F0 (lodrätt) av stavelsen <ma> med fyra toner (hämtad ur Xu 1997:67)

Ingen av tonerna börjar (mellan 0 och 12,5 ms) riktigt rak och jämn som de diakritiska markeringarna antyder, utan med en liten släng som kallas 弯头段 <wāntóuduàn> ordagrant översatt ‘böjt huvud sträcka’, se Wu & Lin (1989:157). Tonen blir stabil efter ca 25 ms vilket kallas 调型段 <diàoxíngduàn> ‘ton mönster sträcka’ och framgår av markeringen ovanför stavelsen. Tonkurvorna förändras vid ca 60 ms och löper sedan vidare till sina respektive slut. Avståndet till slutpunkten kallas 降尾段 <jiàngwěiduàn> ‘fallande svans sträcka’. Termerna motsvaras på engelska i tur och ordning av *onset*, *contour* och *offset* (So & Best 2014).

De olika tonernas kurvor börjar med en viss tonhöjd, se figur 1. Utgående från tonernas onset startar T3 med lägst tonhöjd (knappt över 100 Hz) men även T2 har en låg frekvensstart (något högre än T3) medan T1 och T4 bägge har högre tonhöjd vid starten varvid T4 är den högsta (ca 140 Hz). Tonernas förlopp är olika: T1 ligger nästan konstant kvar omkring 130 Hz, T2 sjunker efter 25 ms men stiger sedan direkt, T3 sjunker kring 50 ms och stiger sedan abrupt medan T4 faller dramatiskt när den vid ungefär 25 ms nått sin högsta frekvenspunkt (nära 140 Hz). Här är kurvornas vändpunkter för T2 respektive T3 av största vikt. De olika egenskaper tonerna har beror på såväl F0-skilnaderna som rörelserna.

Tonerna brukar också beskrivas (Clark et al. 2007:343; Engstrand 2004:193) som nivåtoner (*register*) och konturtoner (*contour*). Med den första termen beskrivs hur högt eller lågt en ton ligger (=frekvensen) medan den andra termen indikerar tonens uppträdande intill dess slut. Lin (2007b:92-96) använder sig av det allmänt vedertagna fonologiska systemet för toner varvid H står för hög tonhöjd och L för låg medan M markerar en ton mellan de två ytterligheterna. En nivåton beskrivs med dubbelmarkering av dess läge, HH eller LL (ibland dock bara enkelmarkering). Konturtonerna får beskrivningar som visar tonens rörelse genom att ange start- och slutpunkter, t.ex. HL eller LH. För kinesiskan blir markeringen för T1 bokstavskombinationen HH, T2 uppträder som MH, T3 indikeras som LH eftersom den börjar L och slutar H medan T4, som är fallande, beskrivs HL. Andra transkriptioner återfinns i Duanmus diskussion (2007:225).

Figuren visar de fyra tonerna. T0 finns inte med, den brukar inte räknas som en egen ton för att dess akustiska egenskaper skiljer sig från de andra tonerna: för det första uppvisar den en mycket otydlig F0 med kort duration och för det andra är den stavelse som den ingår i obe-tonad (Kratochvil 1998; Lee & Zee 2008; Lin 1985; Wang 2004). Begrepp som ton, betoning, accent och deras relationer behandlas i senare avsnitt (2.4).

De ovannämnda akustiska parametrarna, F0, amplitud och duration är viktiga. I kinesiskan, som till sin struktur i princip är ett konsonant-vokalbaserat språk, kan tonerna inte isoleras från segmenten. Här bör påpekas att i den tvådimensionella figuren visas bara tonernas rörelser, F0 och dess duration, och alltså inte samspelet med vare sig amplituden eller de två segmenten [m] och [a]. Figuren visar inte hur den (tonande) initiala konsonanten inverkar på F0 utan är enbart en angivelse för vokalsegmentet. Hittills har vi enbart diskuterat toner i enstaviga yttranden. När en enstavig ton kombineras med en annan ton uppstår dynamiska fenomen.

2.1.2 Tonsandhi akustik

När en ton förändras genom kombination med en annan ton kallas företeelsen tonsandhi. Bruce (2012:134) påstår att även

... om i princip varje (betonad) stavelse har sin egen ton, förekommer det i löpande tal att en ordton inte alltid uppträder i sin originalform utan kan påverkas av och assimileras till angränsande toner, s.k. tonsandhi.

Kinesiskans fyra toner kan tillsammans med den neutrala tonen teoretiskt kombineras (5×5) till 25 möjliga taxotoner, se tabellen nedan.

Tabell 1 Samtliga möjliga tonkombinationer

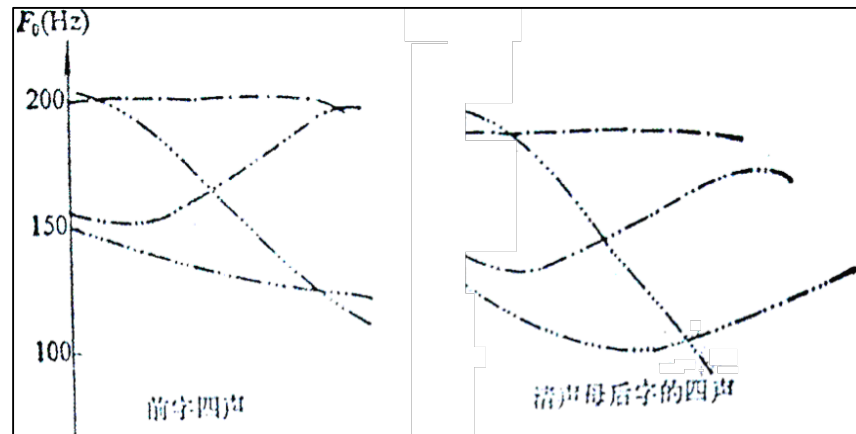
	T1	T2	T3	T4	T0
T1	T1+T1	T1+T2	T1+T3	T1+T4	T1+T0
T2	T2+T1	T2+T2	T2+T3	T2+T4	T2+T0
T3	T3+T1	T3+T2	T3+T3*	T3+T4	T3+T0
T4	T4+T1	T4+T2	T4+T3	T4+T4	T4+T0
T0	T0+T1*	T0+T2*	T0+T3*	T0+T4*	T0+T0*

Den neutrala tonen återfinns aldrig i första stavelsen¹ (markeras med asterisk *) därför återstår 20 kombinationer. Även T3-sandhi (som innebär att T3+T3 ersätts av T2+T3) markeras i tabellen ovan med en * så kinesiskan har i praktiken sammantaget 19 tonkombinationer. Tidigare fonetikforskning, t.ex. Wu & Lin (1989:161-164) visar att i tvåstaviga ord är det tre av tonerna som uppträder på samma sätt som de gör i enstaviga ord: T1 klassificeras som H, T2 som LH och T4 som HL. Enligt dem försvinner "dippen" i T3 i naturligt tal och den manifesterar sig som en låg ton, L. Därmed reduceras T3 till en halv T3 medan alla de andra tonerna behåller sina karakteristika (Xu 1997). Också Lin (2007b:94) säger att "when tone 3 is followed by another tone, it is simply a low tone without the rising part and can be represented as LL", vilket gör att kineser inte förväxlar tonerna med varandra i tvåstaviga ord. Det viktigaste resultatet från Wu & Lin (ibid.) är att den andra stavelsens (σ_2) initialljud har en tydlig inverkan på tonmönstren:

1. när σ_2 har *ett tonlöst ljud* som initial uppstår ett tyst moment direkt efter den första stavelsens slut. σ_2 blir därför tidsmässigt längre än σ_1 .
 - a. Tonerna i σ_1 behåller både frekvens och tonkurva; givet samma ton i bägge stavelserna (t.ex. T1+T1) ligger den som återfinns i σ_2 *lägre* frekvensmässigt.

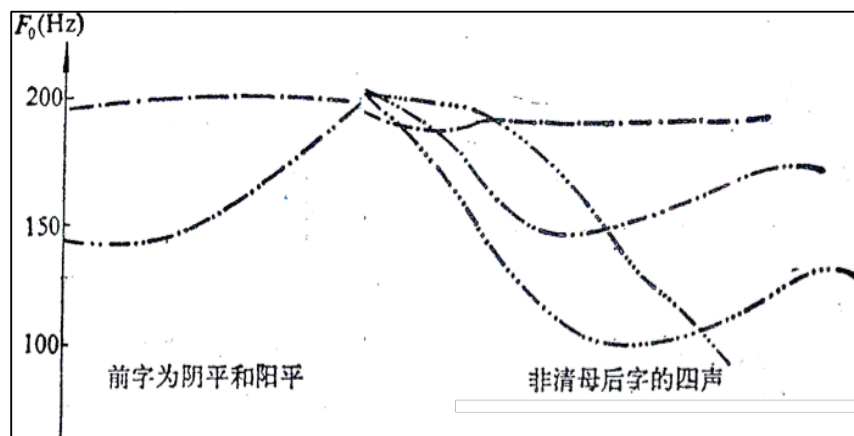
¹ Konjunktivmarkören 的话 <dehuà> är fonologiskt ett tvåstavigt ord där T0 står i σ_1 , men detta ord måste stå i samband med 如果 <rúguō> och dess betydelse blir *om, ifall*

- b. När två T3 kombineras betar sig tonkurvan för den första som en T2 och tonläget är *högre* än T2 i $\sigma 2$.



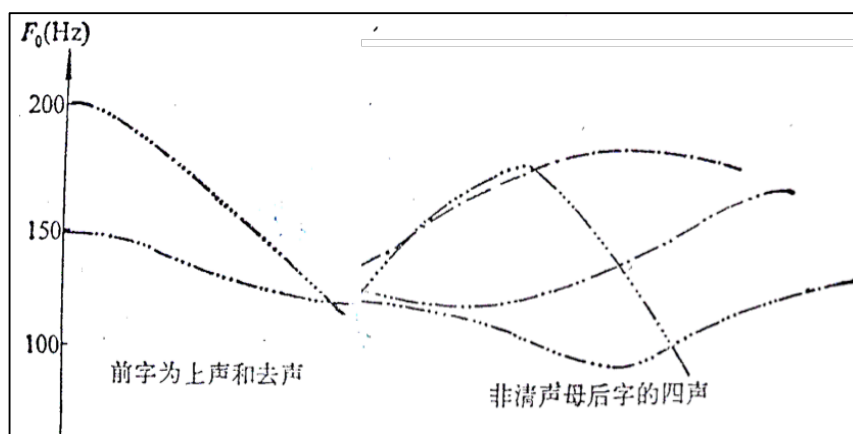
Figur 2 F0-kurvor för $\sigma 1$ (vänster: 前字四声 <qiánzìsìshēng> 'fyra toner för det första tecknet') och $\sigma 2$ (höger: 清声母后字的四声 <qīngshēngmǔhòuzide sishēng> 'fyra toner med ett tonlöst ljud för det andra tecknet'; T1: —, T2: —·—, T3: —··—, T4: —····—) i ledfrasen 这两个字组, 您怎么说? <zhè liǎngge zìzǔ nín zěnmē shuō?> 'Hur uttalar ni det här tvåstaviga ordet X?', en manlig pekingbotalare (hämtad ur Wu & Lin 1989:163)

2. om $\sigma 2$ har en *sonorant initial* sker en övergång av den första tonen till den andra utan avbrott
 - a. om $\sigma 1$ har en T1 eller en T2 (som bägge i sitt slut ligger högt) börjar tonen i $\sigma 2$ på samma höga nivå som föregående T uppvisar i sitt slutskede. Detta leder till att onset för T2 och för T3 sker högre än annars.



Figur 3 F0-kurvor för $\sigma 1$ (vänster: 前字为阴平和阳平 <qiánzì wéi yīnpíng hé yángpíng> 'det första tecknet med T1 och T2') och $\sigma 2$ (höger: 非清声母后字的四声 <fēiqīngshēngmǔ hòuzide sishēng> 'fyra toner med ett tonande ljud för det andra tecknet', T1: —, T2: —·—, T3: —··—, T4: —····—) i ledfrasen 这两个字组, 您怎么说? <zhè liǎngge zìzǔ nín zěnmē shuō?> 'Hur uttalar ni det här tvåstaviga ordet X?', en manlig pekingbotalare (hämtad ur Wu & Lin 1989:163)

- b. om det är en T3 eller T4 i $\sigma 1$ kommer T1 och T4 i $\sigma 2$ att börja på samma frekvens som föregående ton slutade på, d. v. s. på en låg punkt men de strävar till sitt "normala" läge. För T4 innebär detta först en stigning före dess fallande förlopp vilket leder till att den blir förlängd.



Figur 4 F0-kurvor för $\sigma 1$ (vänster: 前字为上声和去声 <qiánzì wéi shàngshēng hé qùshēng> 'det första tecknet med T1 och T2') och $\sigma 2$ (höger: 非清声母后字的四声 <fēiqīngshēngmǔ hòuzide sìshēng> 'fyra toner med ett tonande ljud för det andra tecknet', T1: —, T2: —·—, T3: —··—, T4: —····—) i ledfrasen 这两个字组, 您怎么说? <zhè liǎngge zìzǔ nín zěnmē shuō?> 'Hur uttalar ni det här tvåstaviga ordet X?', en manlig pekingbotallare (hämtad ur Wu & Lin 1989:163)

När ton kombineras med ton handlar det inte enbart om påverkan (där T3-sandhi utgör den extrema följden) utan även om ordbetoning. Om ordets $\sigma 2$ har en T0 framstår den första stavelsen som betonad. Ett minimalt par som 鸭头 <yātóu> 'ankhuvud' jämfört med 丫头 <yātou> 'tös' kan demonstrera företeelsen (märk att i det senare ordet saknas diakritiskt tecken över $\sigma 2$). F0 i den neutrala tonen och dess uppträdande beror mycket på den föregående stavelsens ton (Lee & Zee 2008).

En viss grad av betoning kan upplevas i $\sigma 2$ även när det rör sig om annan ton än T0. Det finns andra minimala par där betydelse ändras enbart genom att ändra tonen antingen i första eller andra stavelsen t.ex. 达人 <dárén> 'intelligent' respektive 打人 <dārén> 'slå någon'. Frågan är vilken stavelse som ska betonas; svaret beror på morfologi, lexikologi och fraseologi. Chao (1968:146-155) och Katrochvil (1998:426-428) antyder att i normala talsituationer tänker talarna troligen inte på de möjliga tonkombinationerna vilket gör att prosodin mera styrs av jamber och trokéer, se även Duanmus diskussion (2007:314f). Ovanstående akustiska fenomen har undersökts i olika tonperceptionsexperiment.

2.2 L1 kinesiska tonperception

I detta avsnitt presenteras först hur kinesiska modersmåltalande uppfattar tonemen (2.2.1) och därefter belyses tonsandhi-perceptionen (2.2.2).

2.2.1 Tonemens perception

Howie (1976) genomförde en serie perceptionsförsök med syntetiserade enstaviga stimuli. Där användes dels syntetiskt tal modellerat efter naturliga F0-mönster, dels syntetiska stavelser där F0-konturerna hade gjorts monotona och dels stimuli som syntetiserats för att låta som viskning (alltså utan stämläppsarbete = ingen F0). Två sorters experiment genomfördes för att bestämma vilket inflytande amplitud respektive duration har på tonidentifikationen. I vardera testet skulle 20 framtagna stavelser tonidentifieras, de presenterades i slumpmässig ordning, fyra teststimuli fem gånger var. Stavelsen *bao* uppfattades korrekt i 95,4% av fallen medan å andra sidan stavelsen *ying* bara nådde ett resultat på 30,8% i sin monotona variant och 39,2% för viskningen; någon förklaring ges inte till valet av stavelser i undersökningen. Howie har vidare diskuterat relationen mellan F0, duration och amplitud och visar att L1-lyssnare känner igen tonerna på ett jämnt och korrekt sätt. Hans slutsats blir att "pitch is the primary feature in the perception of tones" (s. 244).

F0 utgör å ena sidan en väsentlig ledtråd för att identifiera ton (Lin 2007b:92) men å andra sidan rapporterar Guo (1993:80-83) om en studie som genomförts av Liang (1963) som ger

vid handen att viskning identifieras rätt till 50% av L1-talande (högre än Howie 1976, se ovan). I detta försök saknas hänvisningar till kontextens betydelse och till att viskning knappast kan anses vara normal talkommunikation. Yip (2002:292) hävdar att F0 som perceptorisk ledtråd är okontroversiell men betonar att det inte räcker med den för att särskilja en ton från en annan. Andra användbara ledtrådar för tonperceptionen är, som antytts ovan, amplitud (Whalen & Xu 1992) och duration (Chuang et al. 1972; Guo 1993:168-172; Wu & Lin 1989:153-159).

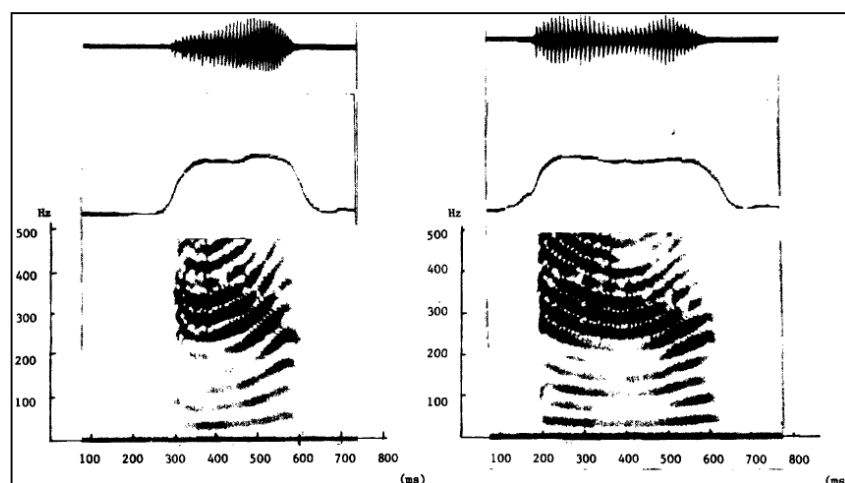
Lin (1988, citerad i Guo 1993:76) och Yang (1986, citerad i Wu & Lin 1989:170) framhåller att amplituden inte spelar någon roll vid tonidentifieringen för L1-kinesiska försökspersoner. Generellt uppvisar T4 relativt stark amplitud jämfört med de andra tonerna och T3 en relativt svag. I motsats mot det föregående visar Whalen och Xu (1992) att deras L1-kinesiska försökspersoner utan några större problem kunde identifiera tonerna utgående från amplituden (frånsett T1, men den är ändå distinkt skild från T3). De tycker att det inte är konstigt att amplituden framstår som oviktig eftersom amplitudkurvorna i Lins undersökning hade redigerats så att de följer tonernas kurvor. Själva tog de bort både F0 och formanterna men behöll durationen före syntetiseringen. I så fall vet man ändå inte om det är duration eller amplitud som inverkar mest på tonidentifikationen. Men Xu påpekar (personlig kontakt, 2015-01-20) att

In experiment 2, duration was controlled. When duration was equalized across the 4 tones, amplitude profiles still enabled tone identification. We didn't test how much duration alone would contribute to tone identification. But it is not hard to imagine how little is left if we have stimuli which have no pitch, no amplitude modulation, but only duration variation. Even if this could lead to some tone identification, for example hearing more T4 if the noise is short, and more T3 if the noise is long, this would not be very surprising or interesting, as far as I can imagine. And I wouldn't be surprised that tone identification with duration alone is much smaller than with amplitude profile.

Av figur 1 framgår att T3 tidsmässigt är längst och T4 kortast så durationen gör det lätt att särskilja dem. Tidsmässigt är T1 och T2 nästan lika långa. Gandour (1983) visar att det som, frånsett duration och amplitud, ger information till dem som har kinesiska som L1 är tonkurvornas uppträdanden. De andra ovannämnda faktorerna kan inte, till skillnad från F0, var för sig avgöra tonidentifikationen, d.v.s. F0 är primär medan de andra faktorerna är sekundära. L1-kinser klarar sig med mindre tydliga markörer (finer-grained cues, Chen 2013a). Skillnaderna kan rentav vara så små som mellan 0.3 och 2 Hz (Klatt 1973).

Då L1-kineser misstar sig på toner sker detta genom att T2 och T3 förväxlas, detta är den slutsats som kan dras av experimenten med syntetiska stimuli där skillnaderna mellan faktorerna var ytterst små. Chuang et al. (1972:299) skriver att "[N]either the falling and rising contour nor the position of the dip point in tone-3 alone can be the perceptual cues to discriminate tone-3 from tone-2". Dessutom visar Li & Thompson (1976) att L1-kinesiska barn kan ha svårt att tillägna sig dessa två toner; deltagarna i deras undersökning lärde sig T1 först, därefter kom T4 och T2, sist kom T3. Förväxlingarna mellan T2 och T3 minskar när det kommer till flerstaviga ord. Gottfried & Suiter (1997) konstaterar att L1-kineser utnyttjar även annan akustisk information för att identifiera ton, t.ex. *knarr* i T3.

Shen & Lin (1991) har visat att det är *the time of the turning point* som utgör ledtråden för skillnaden mellan T2 och T3. För den förstnämnda sker vändningen *före* 400 ms medan den för T3 inträffar *vid eller senare än* 400 ms, jämför figur 5 nedan.



Figur 5 Timing of turning points för <ai> (vänster: T2 före 400 ms respektive höger: T3 efter 400 ms, hämtat ur Shen & Lin 1991:147)

Intressant är att författarna uteslutande har valt kontrastiva ord med de tonande (*voiced*) initiala konsonanterna [l], [m] och [n] och med tystnad förbigår frågan hur/om tonlösa initialer skulle kunna påverka.

Frågan om T2 och T3 utgör skilda kategorier eller de två tonerna egentligen är en och samma kategori är kontroversiell ur psykolingvistisk synpunkt (Zhang 2012), se vidare avsnitt 2.4 och 2.5.

2.2.2 Tonsandhi perception

Lin & Wang (2007a:142-155) syntetiserade åtta tvåstaviga ord, där σ_1 utgjordes av antingen stavelsen <bai> eller <da> med en T1 som konstanthölls på 115 Hz och σ_2 var stavelsen <de> med en syntetiserad T0 vars F0 varierade mellan 70 och 140 Hz. Tio Pekingbor ombads att identifiera tonen i första stavelsen i de nämnda tvåstaviga orden; denna T1 kom att förväxlas på olika sätt alltefter frekvensläget för T0 i den efterföljande stavelsen. Författarna förklarar att försökspersonerna uppfattade stimuli utifrån kontexten. Om de hade svårt att avgöra vilken ton det var i σ_1 jämförde de med den efterföljande tonen och bestämde sig först därefter för vilken ton det var som de hört i σ_1 . De olika uppfattningar om T1 som uppstod angående <bai> och <da> kallas *auditory illusion*.

Duanmu (2007:129-158) hävdar att betoning för kinesiska tvåstaviga ord ligger på första stavelsen. Ett ords prominens är å andra sidan inte viktig för standardkinesisktalande (Wang & Lin 1992:174-182). Shen (2006) skriver att tvåstaviga ord kan upplevas vara betonade beroende av den tonala kontexten. Det är de akustiska faktorerna som är av största vikt för att betoning ska upplevas och då i första hand amplituden medan durationen och tonhöjden inte spelar lika stor roll. Litteraturen om kinesisk ordprosodi ur perceptionsperspektiv är begränsad. Klart är att ett av kinesiskans karakteristika, tonerna, inte är slutgiltigt klarlagt.

2.3 L2 kinesisk tonperception

Forskare har på olika sätt undersökt vilka toner som bereder de lärande störst svårigheter. Redan för 40-talet år sedan genomfördes flera experiment med utländska studenter där det gällde hur de uppfattade kinesiskans toner; försökspersonerna hade skilda modersmål, både icke-tonspråk (oftast engelsktalande) och tonspråk som t.ex. vietnamesiska.

2.3.1 Tonem

Kiriloff (1969) undersökte australiska förstatervisningsstudenter. Det gällde att vid två tillfällen i ett och samma test identifiera 20 stavelser uttalade av en kines med pekingdialekt. Första test-

tillfället var efter två månaders studier och det andra vid terminsslutet och bägge omfattade både transkription och markering av toner. χ^2 -test visade ingen signifikant skillnad mellan de två testen. Han fann att 1) T2 är svårast att uppfatta, 2) T3 tenderar att förväxlas med T2, 3) lägsta antalet förväxlingar inträffade för T4 och T1. Efter andra undersökningstillfällets slut bad han studenterna notera toner utan att de behövde göra någon Pinyin-transkription och nu blev totala antalet korrekta tonnoteringar större vilket leder till tolkningen att segment har inverkan på tonperceptionen. Han föreslår att på nybörjarnivå bör tonidentifiering genomföras åtskild från språkljudsträning men så länge vi inte har kunskap om huruvida segmenten framkallar toner bör detta förslag inte göras allmängiltigt utan får gälla för enbart engelsktalande.

Chuang et al. (1972) spelade upp enstaviga ord (15 med fyra toner) intalade av taiwaneser för fyra L1-kineser och två L1-japaner. Resultaten visar att de flesta felen uppstår mellan T2 och T3. De infödda kineserna missuppfattade dubbelt så ofta T3 för att vara en T2 än vice versa. Japanerna visade inte samma förväxlingsmönster. Forskarna ger inga förslag till vad som orsakar denna skillnad. De har även testat en del tvåstaviga ord, vilket tas upp i avsnitt 2.3.2.

Wang et al. (1999) fann att deras amerikanska informanter hade lika stora svårigheter pretest som posttest att särskilja T2 och T3; distinktionen mellan T1 och T4 befanns vara näst svårast. Ding et al. (2010) skriver att deras tysktalande försökspersoner hade svårast att identifiera T2 och att förväxlingarna där T2 uppfattas som T3 (22%) och T3 till T2 (21%) är de vanligaste. De påpekar att för deras tyska informanter har glottalisering kommit att fungera som identifieringsknep för T3. Liu (2006a) rapporterar att för koreansktalande är tonförväxlingen mellan T2 och T3 vanligast.

Ett annorlunda resultat rapporteras av Liu (2006b) som genomförde ett experiment där 32 vietnameser, uppdelade i två grupper (grund- resp. avancerad nivå), skulle identifiera vissa enstaviga ord. Resultaten visar att för försökspersonerna på grundläggande nivå var T4 svårast att identifiera, medan de avancerade hade svårast att skilja mellan T1 och T4 (i bägge riktningar). Författaren menar att resultaten beror på att också vietnamesiska är ett tonspråk.

Bortsett från perceptionsförsök utifrån rent akustiska experiment visar en annan studie som Hallé et al. (2004) genomfört att tonidentifikationerna för deras taiwanesiska försökspersoner mer låg på ett fonologiskt plan (*quasi-category*) till skillnad från de fransktalande som i högre utsträckning utgick från de akustiska skillnaderna (*psychophysical*). Avsnittet om andraspråksinlärningssteori (Second Language Acquisition, SLA) berör området vidare.

2.3.2 Tonsandhi perception

Guo (1993:330-333) visar att kvoten felaktigt uppfattade toner ökar med antalet stavelser (från en till fyra) i ordet. Liu et al. (1989) testade en-, två- och trestaviga ord och visade att förväxlingarna blir flera allteftersom antalet stavelser ökar.

I det tidigare nämnda experiment som Chuang et al. (1972) genomförde presenterades även tvåstaviga ord, 16 tonkombinationer uppdelade i fyra betydelselösa och 64 betydelsebärande ord. Det visade sig att förväxlingarna mellan T2 och T3 i $\sigma 1$ blev mycket färre än i $\sigma 2$ och att de betydelselösa orden missuppfattades oftare än ord med innebörd. I forskningen om tonkombinationer har Chuang et al. testat tonsandhi i tvåstaviga ord med förväxlingen från T3+T3 till T2+T3; deras slutsats är att det är oklart om det är detta fenomen som utgör hela förklaringen till feltolkningarna mellan T2 och T3.

Wang (1998) utförde ett perceptoriskt experiment omfattande 64 tvåstaviga ord med 16 tonkombinationer (ingen kombination med T0 fanns med) som skulle avlyssnas av 25 studenter med ungerska som L1. Försökspersonerna hade fått orden skrivna med Pinyin och skulle enbart markera tonerna. Resultaten visar att $\sigma 1$ utsätts för fler felaktiga markeringar än $\sigma 2$. Tonkombinationerna T1+T1, T4+T4, T4+T1 och T1+T3 var lättare att identifiera än T3+T1, T3+T3, T4+T2 och T2+T2. Rapporten visar en tydlig tendens för att vissa tonkombinationer dras till enbart en eller två felresponser. Som exempel kan nämnas att de flesta informanterna

felmarkerade tonkombinationen T1+T2 antingen som T1+T3 eller som T4+T2. Det ges ingen förklaring till varför dessa förväxlingar skedde.

I en del av sitt examensarbete rapporterar Liu (2006a) att för koreanska studenter var T1+T3 den stimulus som var svårast att uppfatta när de lyssnade på 32 tvåstaviga ord. När det gäller de vietnamesiska studenterna gjorde de, enligt Liu (2006b), största antalet felmarkeringar för T4+T4, vilket tros bero på L1-transfer (se vidare under avsnitt 2.5).

Ding (2012) redogör för ett perceptionsförsök med tvåstaviga ord, både sådana som var kända och okända för lyssnarna. När det gäller första stavelsen identifieras T4 och T3 bättre än T1 och T2. Det antas att misstagen mellan T1 och T2 beror på att dessas toner slutar på likartad tonhöjd och samma antagande görs för förväxlingarna mellan T3 och T4. Förväxlingarna mellan T2 och T3 tolkar man bero på att startpunkten i tonhöjden ligger tämligen lika. För σ_2 rapporteras att det var T2 och T3 som fick flest felmarkeringar. Misstagen mellan T2 och T3 skylls även de på likheten i tonhöjd vid tonernas startpunkter medan förväxlingarna mellan T3 och T4 tros bero på den likartade frekvensen i tonernas slutpunkter. Ding påpekar i artikeln att "den professionelle inläsare" som de använde har en T3 som avviker från vad som räknas vara rätt men påpekar att detta "can be found in many standard speakers, and there is no glottalization in the middle of his third tone". Ding et al. (2010) problematiserar inte den identifieringsskillnad för T3 mellan enstaviga respektive tvåstaviga ord som framgår av deras rapport. För de enstaviga anges (som nämnts tidigare) knarret vara ett identifieringsknep medan de själva explicit betonar att sådant inte förekommer när T3 återfinns i den andra stavelsen i tvåstaviga ord. Som nämnts tidigare (avsnitt 2.1.2) försvinner den dipp som T3 gör i löpande tal.

De vanligaste tonemförväxlingarna sker mellan T2 och T3 och detta oavsett om informanterna är L1-kineser eller personer med annat modersmål. Även när vi tittar på tvåstaviga ord verkar, oavsett L1, problemet ligga i T2 eller T3, så t.ex. för L1-ungerska som uppfattade T1+T2 som T1+T3 och stimulus T2 i andra stavelsen uppfattades som T3. Det mest vanliga är att forskarna konstaterar felmarkeringarna och sina informanternas tillkortakommanden men sällan får vi förklaringsförsök till dessa.

Engelsktalandes kunskaper om det engelska intonationssystemet assimileras förmodligen till och tillämpas på tonerna i mandarin. Särskilt lätt att uppfatta var den fallande T4 när den uppträdde finalt i ett yttrande, en position där en liknande kontur är att förvänta i engelskan (Broselow et al. 1987). När det gäller T1 och T4 kan detta kanske bero på att engelsktalande lyssnare uppfattar kinesiskans höga toner (T1 och T4) som betoningar medan den låga T3 uppfattas som obetonad, allt detta trots det faktum att stavelsebetoning (enstaviga ord) i mandarin snarast manifesteras som skillnader i duration och amplitud än i F0-skillnader (White 1981). Shen (1989) förklarar att T1 och T4 är svåra att ta till sig för amerikaner och hävdar att detta beror på L1-engelskastörning och skriver vidare att T1 och T4 är mer utsatta för påverkan från L1 eftersom de inte är lika prosodiskt markerade som T2 och T3. Kritiskt bör framhållas att hennes studie är baserad på en talproduktionsanalys av L1-amerikanskt uttal av L2-kinesiska.

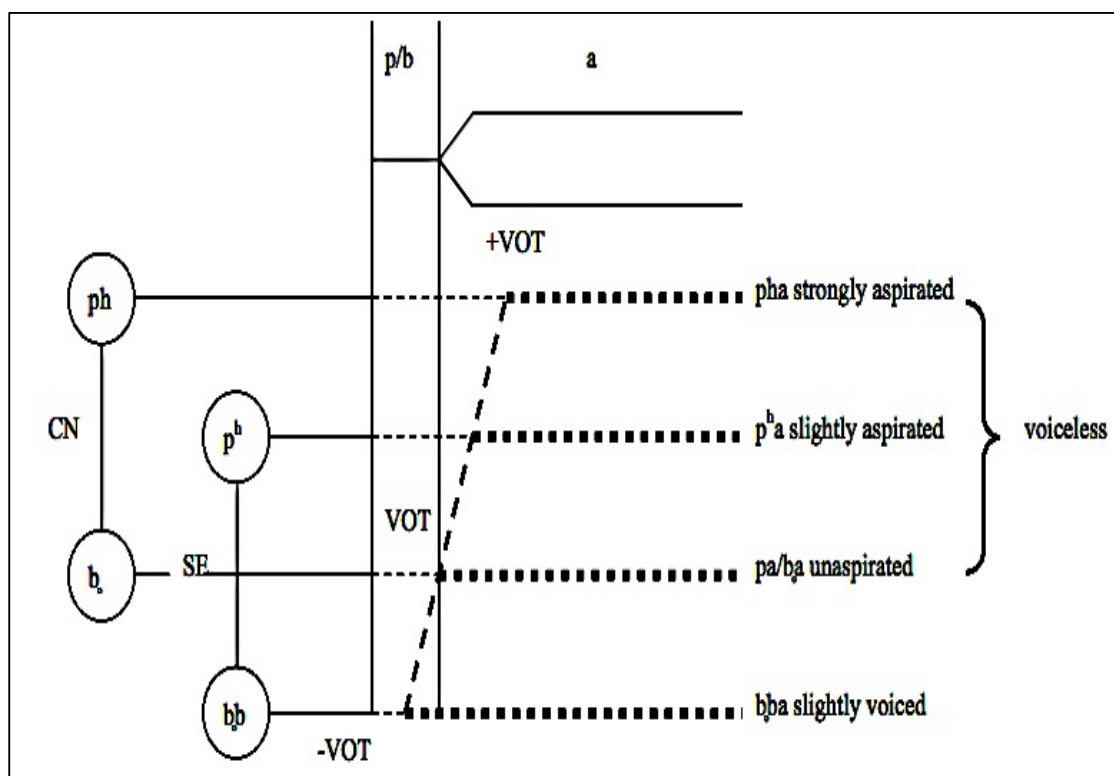
Gottfried & Suiter (1997) undersökte om L1-kineser använder sig av knarr i T3 (så även Ding et al. 2010). Gårding et al. (1986) godtar inte knarret som avgörande faktor när det gäller att skilja T3 från T4. Enligt dem uppmärksammar L1-kineser den stigande onsets för T4 och dess fallande motsvarighet i T3 men forskarna utesluter inte att knarret kan vara en del av diskriminationen mellan å ena sidan T1 och T2 och å andra sidan T3. När de L1-tyskttalande i Dings studie lyssnade till tvåstaviga ord uppmärksammades inget knarr.

Klart är att det inte är tillfyllest att analysera L2-perception utifrån L1 kinesisk akustik i enstaviga ord (i termer av F0, amplitud, turning point, knarr och val av testord). Hänsyn bör tas till den lyssnandes egen ordprosodi i flerstaviga ord. Vad mer än akustisk-perceptoriska experiment krävs för att vi ska få en större bredd på bilden av tonmissuppfattningarna gällande L2? Är skillnaderna generella eller krävs det specifika jämförelser språkvis för alla

språk kontra kinesiska? Är sådana aspektskillnader att söka på fonetisk eller fonologisk nivå? Sammanfattningsvis kanske man skulle fråga sig om tonperceptionen beror på vad talaren åstadkommit eller inverkan från lyssnarens modersmål.

2.4 Jämförelse mellan kinesiska och svenska

Kinesiskan skiljer sig från svenskan både vad det gäller ljud- och tonsystem. F0-varianter förekommer i bägge språken men aspirationssystemen är olika. De kinesiska konsonanterna delas in i två huvudsakliga kategorier, sonoranter [+son] och obstruenter [-son]. Sonorantkategorin består av approximanter [+app] och nasaler [-app]. Approximanterna, slutligen, återfinns som lateraler [+lat] och centraler [-lat]. Obstruenterna omfattar två typer: klusiler [+klu] och frikativor [-klu]; klusilerna i sin tur indelas i [+asp] och [-asp]. Bannert (enligt sammanfattning av Garlén 1988:20-30) typologiserar de svenska klusilerna som aspirerade och oaspirerade och menar att detta skulle stämma överens med mandarin. Helgason & Ringen (2008) presenterar en detaljerad akustisk undersökning av svenska aspirerade och tonande klusiler. Skillnaden [$\pm$ tonande] rörande klusilerna existerar inte i kinesiskan som istället uppvisar varianterna [$\pm$ aspirerad], se figur 6.



Figur 6 Jämförelse mellan svenska och kinesiska klusilerna enligt Catford's modell (hämtad ur Hu 2012)

Figuren visar skillnaden mellan klusilerna i bägge språken enligt Catfords modell. Av figuren kan vi se en stor tidsskillnad mellan klusilerna. Översikter för dessa kinesiska klusiler jämfört med andra språk återfinns bl.a. hos Catford (2001:181-186) och Lin (2007b:41-43). Skillnaden mellan klusilerna i svenska och kinesiska är fonologisk och kräver akustisk verifikation.

Det här systemet med (o)aspirationer uppmärksammades i både praktisk pedagogisk verksamhet och teori. Hu märkte att en del svenskar som efter en tids undervisning perceptoriskt klarade av de kinesiska tonerna fortfarande hade svårigheter att åstadkomma dem i kombination med vissa ljudkategorier. Studenterna upplevde att dessa ljud stör tonperceptionen och dels att deras muskler i talapparaten var ovana att säkert producera dessa kombinationer.

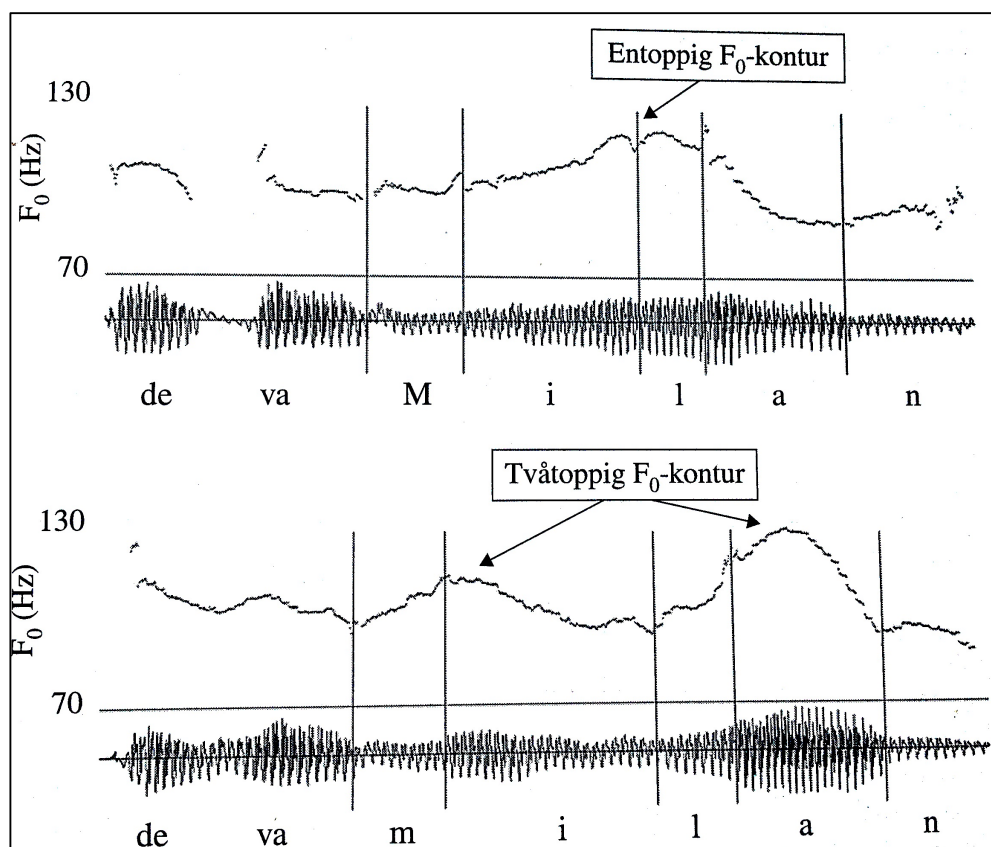
Vissa av kinesiskans fonologiska särdrag verkar vara speciellt svåra för svenskar, t.ex. uppfattas kinesiskt [-asp] som svenskt [+ton] medan å andra sidan ett kinesiskt [+asp] tolkas som [-asp]/[-ton]. De berättade även att det verkade vara den annorlunda kinesiska aspirationen som påverkar deras tonperception.

Hombert (1978) studerade engelska, franska och yoruba varvid tonhöjden befanns vara högre efter tonlösa klusiler än efter tonande. House (1990) skriver att det verkar finnas gränser för hur toner kan uppfattas och då speciellt för timing, både vad gäller perception och uttal. Han menar att känslan för tonhöjden inte är konstant utan ändras med komplexiteten i talsignalen. Hørseln verkar inte klara av att ta hänsyn till både spektral och tonal information samtidigt. Till exempel återfinns ett moment med maximal spektral förändring när ett konsonantljud övergår till vokal (som i betonad stavelse), så känslan för tonhöjd minskar och tonrörelsen kan fokuseras. Mittpartiet av en vokal uppvisar en större spektral stabilitet som ökar känsligheten för tonhöjd eftersom perceptionen inte tvingas ägna sig åt att identifiera tonrörelsen. Nästa övergång mellan vokal och konsonant uppvisar ett nytt steg i spektral förändring vilket åstadkommer att tonhöjds-känsligheten sänks momentant. Sammanfattningsvis innebär detta att det blir en ojämn fördelning av tonhöjdsförändringar i relation till olika delar i en betonad CVC-stavelse (Consonant-Vowel-Consonant struktur som i ordet <lat>).

Tonperception bör bredda sig från åsikten att det handlar bara om F0 som är så lätt att upptäcka i vokalen. Kinesiska är ett typiskt CV-språk där tonen gör sig gällande redan före vokalen. Xu & Xu (2003) har visat att F0-stabiliteten påverkas starkt av aspirerade obstruenter i kinesiskan. Akustiskt leder aspiration till att F0 sänks medan en frånvaro av aspiration leder till en höjning - men detta stämmer inte överens med vad Hombert (1978) anser, denna diskrepans kan bero på att de har olika utgångspunkter, ett akustiskt respektive ett perceptoriskt experiment och språk med olika fonetiska uppsättningar. Kvar står att olika konsonanter påverkar F0.

När vi vänder oss till tonsystem vet vi enligt ovannämnda baskunskaper i kinesiska att där finns fyra toner och 19 tonkombinationer. Det är fortfarande en kontroversiell fråga om hur många underliggande tonkombinationer det möjligen kan finnas när även begreppet betoning räknas in. Svenska uppvisar både ett betonings- och tonaccentsystem till skillnad mot kinesiska där tonsystemet tillåtit överskugga allt annat. “[O]rd som består av enstaviga stammorfer, t.ex. bil, är oproblematiska – betoningen ligger självfallet på den enda tillgängliga stavelsen” (Engstrand 2004:178). Dessutom har “[E]nstaviga ord alltid akut accent” (ibid. s.189). På enstavig nivå har kinesiskan fyra toner (svenskan bara en) och de stavelser som har ton är också betonade.

Vad beträffar tvåstaviga ord har den grava accenten “en tvåtoppig F0-kontur som markerar att två eller flera stavelser hör till samma ord” (ibid.). Dessutom har sammansättningar en huvudbetoning och en bibetoning (den grava accenten) medan däremot två enskilda ord uppvisar varsin akut accent; morfologin styr detta i stor utsträckning (Gårding 1998).



Figur 7 Vågformer och F₀-kurvor för orden Milan (entoppig) och milan (tvåtoppig) i ledfrasen Det var _ Manlig talare (hämtad ur Engstrand 2004:187).

Den akustiska skillnaden ligger tydligen i grundtonsförloppen: F₀-kurvorna i figuren är olika och förhåller sig på olika sätt till den segmentala strukturen. I "Milan" befinner sig F₀ relativt lågt vid början av den betonade vokalen för att sedan stiga. Därefter sjunker konturen under den andra vokalen. På så vis bildas en F₀-topp (akut accent) ungefär vid mitten av ordet, vilket verkar likna T2+T4. Jämför "milan" där F₀ stiger mot ett första maximum vid början av den betonade vokalen, sjunker under denna vokal för att åter stiga till ett andra maximum i den sista vokalen. Vi får en tvåtoppig F₀-kontur, grav accent, som liknar den kinesiska tonkombinationen T4+T2. Bruce (2005) bekräftar att tidigare beskrevs stockholmsmålets Accent I i isolerad ställning som ett rise-fall-mönster och Accent II som fall-rise-fall (två tonhöjdstoppar) med betoningen på första stavelsen.

Intressant är att forskarna har konstaterat tontopparna men också lagt märke till interaktionen mellan den segmentala strukturen och F₀. Konstateras kan att vår kunskap fortfarande inte har kunnat ge en slutgiltig kunskap – skillnader mellan de två språken finns dels vad avser F₀ och dels tidsskillnaderna mellan F₀s uppträdande. Skiljer sig perceptionen från akustiken trots likheterna? Sker en perceptorisk överföring och/eller assimilering till svensk tonaccent av kinesiska tvåstaviga ord? Skulle vi kunna förutsäga något med hjälp av några klassiska teorier inom anspråksinlärnings teorier (Second Language Acquisition, SLA)?

2.5 SLA-teori och kinesisk tonperception

(C)ontrastive (A)nalys (H)ypothesis, CAH, och transferteorin (Abrahamsson 2004:79-116; Abrahamsson 2009:235-253; Gass & Selinker 2008:95-102; Lightbown & Spada 2006:77-108) innebär kortfattat att inläring sker utifrån kategorisering av de likheter och skillnader man upplever från abstrakta och konkreta mönster. Detta innebär att svårigheter lättast upp-

står när/om mönstren uppvisar bara minimala skillnader i form eller betydelse. Med utgångspunkt i det ovanstående borde det vara lätt för svenskarna att uppfatta kombinationerna T2+T4 och T4+T2. Enligt den starka versionen av hypotesen, den s.k. a priori-versionen, kan man med hjälp av en kontrastiv analys – en systematisk jämförelse – av två språk predicera de svårigheter och fel som kommer att dyka upp. Vid andraspråksinläring används ibland strukturer/regler från t.ex. svenskan när dessa överensstämmer med kinesiskan, men om strukturerna i de två språken är olika kan felaktigheter uppstå. Någon systematisk jämförelse vad gäller ordprosodin i de två språken har ännu inte publicerats och inte heller finns det någon vetenskaplig forskning angående assimilation av ljuden i de två språken.

Förutom CAH och transfer har det tagits fram modeller för att förklara de svårigheter som vuxna har vad gäller perceptionen när de lär sig ett L2, bl.a. Speech Learning Model (SLM, Flege 1995:233-277) och Perception Assimilation Model (Best 1995:171-206). Bägge är inriktade på segmentnivå men PAM har senare utvecklats med PAM-Suprasegment (So & Best 2014).

SLM fokuserar på sambandet mellan perception och produktion vid inlärandet av ett annat språk. Tre faktorer listas som försvårar förvärvet av det nya språket: dels i vad mån modersmålet fortsatt används i hög utsträckning, dels de annorlunda kontexterna för språken och dels det sätt som det nya språket erbjuds. SLM betonar att svårigheterna kan ske i båda riktningar och att det är en process som inbegriper både perception och produktion. SLM kan även fungera i motsatt riktning, dvs. att ett gemensamt fonetiskt utrymme delas upp i två kategorier. Båda processerna kan resultera i kategorier som inte finns hos L1-talarna och leda till resultat som varierar från bra till avvikande på behärska-skalan.

PAM har framlagts som förklaring till varför främmande fonem i ett L2 är så svåra för vuxna inlärare. L2-fonem kan enligt modellen assimileras på två sätt, antingen som språkljud eller icke-språkljud. I fråga om språkljud kan ett L2-fonem överensstämma med kategorier i det egna L1 och då föreligger det inga problem men det kan också finnas gradskillnader alltifrån överensstämmande till märkbart avvikande. L2-fonem kan också bli okategoriserade och upplevs då sakna anknytning i L1.

När segment i det nya språket liknar kategorier i L1 assimileras de (PAM, Best 1995) eller klassas som likvärdiga (SLM, Flege 1987) med dessa; här finns likheter mellan de två modellerna. L1-påverkan är inte begränsad till fonologisk kontrast, PAM betonar uttryckligen att de fina ljudskillnaderna/-likheterna mellan L1 och L2 är viktiga. SLM framhåller att relationen mellan ljuden i L1 och L2 uppfattas i relation till varandra snarare än på den mer abstrakta fonemiska nivån; kategorierna i de två språken återfinns i ett gemensamt fonologiskt utrymme. Interaktionen språken emellan kan utvecklas till en assimilation eller sammansmältning av L1- och L2-egenskaper så att L2-studerande tror sig höra något som egentligen utgör ett mellanting mellan L1 och L2. Ett närliggande exempel utgörs av att svenska studenter uppfattar kinesisk [±asp] som L1-svensk [–ton].

So & Best (2014) genomförde ett experiment rörande infödda kinesers, fransmäns och engelsmäns tonperception. Fransmännen och engelsmännen hade inte studerat kinesiska och inte heller fått någon musikalisk utbildning (jfr. Bidelman et al. 2011, nedan). De fick lyssna till “zhi, zha, zhu, fu” uttalade av en L1-kinesisktalande i meningen “xia4 yi1 ge4 shi4 X zi4” *the next one is the X word* i påstående, frågande och utropande sats. Både fransmännen och de engelsktalande kategoriserade kinesiska toner utifrån fonetiska likheter med intonationskategorier i sina egna L1. Fransmännen visade sig bättre kunna urskilja de små fonetiska skillnaderna mellan T3 och T4, helt i överensstämmelse med betoningsmönstren i sitt modersmål. Trots detta var missuppfattningarna mellan T1 och T4 respektive T2 och T3 vanligare än för andra förväxlingar. Författarna konstaterar att förväxlingarna T1/T4 och å andra sidan T2/T3 är kontrastiva och stämmer överens med Hallé et al. (2004, så även Ding et al. 2010); de verkar också verifiera vad redan Karlgren (1965:30), utan tillgång till dagens avancerade utrustning, kunde konstatera:

jämn ton (ungefär som på svenska på ett lugnt konstaterande *ja* ⁀);
rask och rakt stigande ton (t.ex. svar på ett tilltal: *já* = vad vill du?);
långsamt och brutet stigande ton (t.ex. ett tvekande *jåå*);
fallande ton (t.ex. ett glatt, triumferande *jå!!*=där ser du!)

Det kan alltså framstå som om PAM-S faktiskt kan förklara hur L1-intonation assimileras till L2-kinesiska tonsystem. Om kinesiska språket inte hade utvecklats sedan Karlgrens epok skulle denna modell vara utmärkt. Men som tidigare nämnts består modern kinesiska inte enbart av dessa enstaviga ord utan också flerstaviga ord. Därför kan det inte vara rimligt att kringgå att det existerar kinesisk ordprosodi. Detta är den första av två svagheter i de presenterade modellerna som vid en första anblick verkar vara bra verktyg för att testa hur olika L1-ljud respektive -toner assimileras linjärt till L2-ljud respektive -intonation. Men talsignaler är inte isolerade enheter, uppdelade var för sig som segment och suprasegment, de är tätt sammanvävda med varandra. Särskilt komplicerad är den interaktion som äger rum över stavelsegränserna (Bannert 1994:25). Hur segment påverkar suprasegment kan modellerna inte förutsäga, och detta är deras andra svaghet.

SLM betonar AOA (Age of Arrival) mer medan PAM fokuserar mest på om personerna inte har någon kunskap i L2 (naiva). Frågan är hur mycket dessa modeller skulle kunna förklara eller vara tillämpbara för vuxna som studerar ett L2 på nybörjarnivå. Tidigare la forskarna tonvikt vid att utforma testen med utgångspunkt från sina hypoteser så det kan ifrågasättas på vilket sätt dessa resultat reflekterar egentliga problem för L2-kinesisktalande i perspektivet av språkinläring och -undervisning. Sällan har det samlats in data direkt från studenter. Vad i L2 är det som påverkar perceptionen och vad i L1 leder till perceptuell brytning (McAllister 1996; jfr. foreign accent i Strange 1995:22)? Ingen av modellerna verkar lägga tonvikt (positivt eller negativt) på den ömsesidiga påverkan som sker på vuxnas nybörjarnivå.

För vuxna sker jämförelser i högre grad med de kategorier som modersmålet innehåller, det sker en assimilation av det som är nytt till de kategorier som redan finns i L1. Följden blir att det kan uppstå problem när de lyssnar till icke-infött tal. Detta är modellerna överens om trots att deras teoretiska underlag och fokus varierar. Parentetiskt kan påpekas att svenskar, före sina studier i kinesiska, oftast har undervisats i minst ett främmande språk och troligtvis tar med sig erfarenheter/regler/kategorier från dessa studier in i det nya språket.

2.6 Sammanfattning

Kontrollerade akustiska experiment ur L1-kinesiska perspektiv tyder dels på att det är F0 som indikerar den ton som skiljer enstaviga ords betydelser, dels att duration och amplitud spelar in. När ton kombineras med ton kommer deras akustiska egenskaper att ändras mer eller mindre p.g.a. tonkontexten och dessutom inverkar t.ex. betoning och andra prosodem. Litteraturläsningsgenomgången visar att forskningen om L2-kinesisk tonperception är kontroversiell beroende på de inlärandes olika L1-bakgrunder, vidare att testmaterialet varierar och även på att det finns skilda utgångspunkter för forskningen avseende tonmissuppfattningar. I all denna diversifiering finns ändå resultat som i viss mån är gemensamma, t.ex. att tonförväxlingen mellan T2 och T3 är den vanligaste (och förekommer ibland t.o.m. hos L1-kineser). Någon fullständig förklaring kan ännu inte presenteras enligt Sebastián-Gallés (2005) som har gjort en översikt över de metoder som använts för att beforska prosodi.

Sällan har det nämnts något om att segment skulle kunna ha inverkan på suprasegmenten (tonerna). Förmodligen spelar detta inte någon roll för L1-kinesisktalande men det borde kunna inverka på L2-kinesisk tonperception. Frånvaron av betoning kan leda till att en tons utseende blir mindre accentuerat, ett fenomen som behöver lyftas fram inom tvärspråksforskningen. Hur sker perceptorisk assimilering av svensk ordprosodi (accent och betoning) till kinesiska? Här skall återigen betonas att det inte är enstaviga ord det handlar om utan minst tvåstaviga och att det då finns både ton och betoning i kinesiskan. Här uppstår frågan om det

perceptoriskt sker en överföring och/eller assimilering till svensk tonaccent av kinesiska tvåstaviga ord.

Ovanstående litteraturgenomgång visar att tonperception är en komplicerad och dynamisk process. Sådant kan vare sig PAM eller SLM förutsäga för dessa modeller gäller enbart på linjär (supra-)segmentell nivå.

3 Frågeställning

Det som undersökningen vill reda ut är om svenskar har samma problem som andra L2-studerande, nämligen: 1) om T2 är svårast att identifiera, 2) om tonförväxlingen mellan T2 och T3 är vanligast; 3) om det sker tonidentifikation och tonförväxlingar på samma sätt mellan σ_1 och σ_2 ; 4) om konsonanttyperna påverkar tonperceptionen p.g.a. att ljudsystemen skiljer sig åt mellan språken och i så fall vilken konsonantkategori som inverkar mest och 5) om någon tonkombination i högre/lägre grad anknyter till det svenska tonaccentsystemet. Sammanfattningsvis: använder svenska lyssnare företeelser i sitt eget språk för att få grepp om strukturerna för ordprosodin i kinesiskan? Med dessa funderingar genomfördes det experiment som presenteras nedan.

4 Metod

Det är omöjligt att testa alla tidigare forskningstester, dels p.g.a. att kursplan och tidsplanering inte tillåter det, och dels för att vi inte vet vilka andra L1-bakgrundsföreteelser som svenskar som studerar kinesiska bär med sig. Detta är bakgrunden till att de data som analyserats för detta arbete utgörs av inlämnade svar på sluttentamen efter studenternas första termin. Stimuli är autentiska och inte syntetiska eftersom det är orealistiskt att använda syntetiska ljud i den löpande undervisningen. Det insamlade materialet utgörs av diktamen (tvåstaviga ord) under de tre akademiska åren 2009 t.o.m. 2011 så egentligen kan det sägas utgöra tre resultat. Tentamensproceduren var likadan varje gång. Deltagarna belönades inte finansiellt på något sätt för deltagandet i provet eftersom examinationen var obligatorisk.

4.1 Informanter

Alla 94 deltagarna (54 män, 40 kvinnor) var mellan 18 och 25 år och vid studiens början ny-registrerade heltidsstuderande på nybörjarkursen i kinesiska på Institutionen för språk och litteraturer, Göteborgs Universitet. Ingen av dem uppgav sig vara hörselskadad, inte heller hade någon tränat musik professionellt². Mer data om försökspersonerna återfinns i tabell 2.

Tabell 2 Uppgifter om informanterna

Kön	Läsår-09/10		Läsår-10/11		Läsår-11/12		Totalt	
	L1-Sv	L1-A	L1-Sv	L1-A	L1-Sv	L1-A	L1-Sv	L1-A
Man	13	3	16	3	16	3	45	9
Kvinna	5	6	12	2	9	6	26	14
Totalt	18	9	28	5	25	9	71	23
		27		33		34		94

Horisontellt visar tabellen läsåren. Markeringen L1-Sv respektive L1-A anger att försökspersonen har svenska eller annat språk som modersmål. 71 (45 män, 26 kvinnor) hade svenska som modersmål och saknade förkunskaper i kinesiska. De hade studerat ca 3,5 månader på

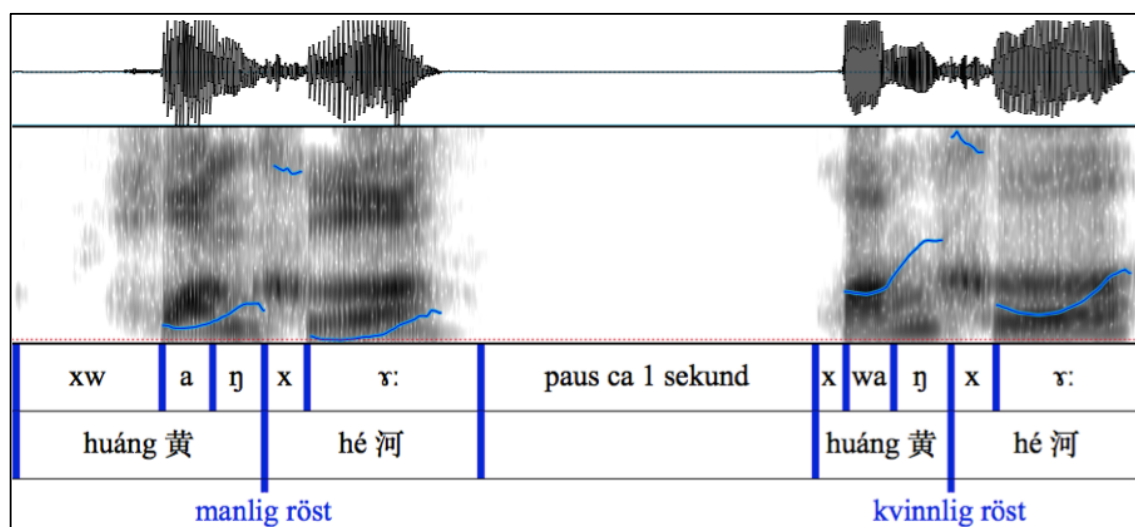
² Bidelman, Gavin M, Jackson T Gandour & Ananthanarayan Krishnan (2011). Cross-domain effects of music and language experience on the representation of pitch in the human auditory brainstem. *Journal of Cognitive Neuroscience* **23**, 425-434. De fann att L1-kineser och L1-engelska amatörmusiker hade betydligt bättre uppfattning angående kinesiska toner än engelsmän utan musikalisk träning.

institutionen efter registreringen. 5 deltagare (2 män, 3 kvinnor) hade skånsk dialekt medan de övriga hade västra-götalandsmål som dialekt. 23 av det totala antalet deltagare hade annat modersmål än svenska: 11 (4 män, 7 kvinnor) var svensk-kineser med olika kinesiska dialekter (t.ex. kantonesiska och min), två hade vare sig svenska eller kinesiska som modersmål utan finska (kvinna) respektive ryska (man). 6 studerande (3 av vardera könet) hade läst kinesiska i olika utbildningsformer i Kina under en tid av från fyra månader upp till två år. Fyra (1 man och 3 kvinnor) hade läst kinesiska som fristående kurs på gymnasienivå.

4.2 Stimuli

Två studenter från Kina, båda 24 år (en man och en kvinna), som 2008 bodde i Göteborg, spelade in testmaterialet på frivillig basis och utan ersättning. Den förstnämnde studerade på Chalmers och den andra vid Göteborgs universitet. De uppgav att deras föräldrar och generationen före dem hela tiden varit bosatta och levit i Beijing. Deras uttal var så tydligt och välartikulerat att examinator och personer med kinesiska som modersmål godtog deras inläsning som stimuli i tentamen.

De två inläsarna spelade in testorden på Labb 1 på Humanistiska Fakulteten vid Göteborgs Universitet. En mikrofon med hörlurar, modell Tandberg Educational, Total Impedance 200 Ohm, användes. Avståndet mellan mikrofon och mun var ca 12 cm för att undvika störande ljud som t.ex. aspiration. Labbet var tyst. Varje ord lästes tre gånger, och spelades in i format 16 bitar upplösning och samplingsfrekvensen 44,1 kHz i ljudanalysprogrammet Praat, se figur 8 nedan.



Figur 8 Vågform (överst), spektrogram (mellan) och blåmarkerat F0-kontur av ett test ord nr. 1-黄河-<huánghé>-[xwaŋA-xɿA] 'Gula Floden'. Manlig respektive kvinnlig talare

Därefter sparades ljudfilen i wav-format och redigerades på så sätt att den andra inlästa versionen av ordet kom att användas i studien eftersom den lät mest naturlig. Intervalltiden mellan den manliga och kvinnliga rösten begränsades till ca 1 s. Även tidsintervallet mellan orden är ca 1 sekund. Figuren visar den manliga (till vänster) respektive den kvinnliga rösten (till höger). I ordning uppfifrån och ner återfinns oscillogram, spektrogram och TextGrid (IPA, Pinyin och tecken). Rösternas tonkurvor är angivna (blå heldragen linje).

Ljudmaterialet omfattade 25 kinesiska tvåstaviga ord som hade hämtats från de icke-behandlade kapitlen 15 – 17 i kursboken (T'Ung & Pollard 1983). När terminen avslutades hade kursen under alla tre åren nått fram till kapitel 9. De 25 orden valdes av examinator, en numera pensionerad svensk lärare med lång erfarenhet av undervisning i kinesiska. Orden sorterades i slumpmässig ordning, se bilaga 1, och kategoriserades i fem kategorier utifrån det

kinesiska konsonantsystemet (jfr 2.4, se även bilaga 2): 1) lateraler [+son, +app, +lat] (förkortad [+lat]), 2) centraler [+son, +app, -lat] ([-lat]), 3) aspirerade klusiler [-son, +klu, +asp] ([+asp]), 4) oaspirerade klusiler [-son, +klu, -asp] ([-asp]) och 5) frikativor [-son, -klu] ([-klu]), se detaljerad förklaring i bilaga 2. Fördelningen är ojämn: alla finns inte representerade bland testorden, så t.ex. finns ingen initialnasal medtagen. Materialet täcker inte på långt när alla kinesiska CV-sturkturer (ca 400 utan toninverkan respektive 1300 medräknat toner, Wu & Lin 1989:194). Inga testord har i sin helhet samma uttal, däremot förekommer homografi, homofoni och homotonalitet, se bilaga 1.

4.3 Procedur

Under tentamen satt deltagarna i labbet. Miljön var tyst. De använde samma modell av hörlurar med mikrofön (Tandberg Educational, Total Impedance 200 Ohm) som inläsarna hade haft. Svarsformulären var anonyma och försedda med ett identifikationsnummer som samordnats av examinationsavdelningen. Försökspersonerna instruerades att dels med Pinyin skriva ner vilka stavelser och dels ange vilka toner de tyckte sig uppfatta. De fick lyssna på materialet så många gånger de kände sig behöva, det gällde att undvika den tidspress som utgör del av vissa psykolingvistiska experiment (t.ex. reaktionstider i Lee et al. 2010). I experimentet tillämpades inte heller tekniken att "underline the word you hear: Tik; Tiyk; Tis; Tiys" (Modern Language Aptitude Test, MLAR, Carroll & Sapon 1959, citerad i Gass & Selinker 2008:419), detta för att vi inte ville begränsa informanternas svar utan hålla öppet för så många felaktiga varianter som möjligt. En annan anledning till att det krävdes av informanterna att stava var att det finns experimentdata som visar att "the processing of tonal information may be more error-prone than the processing of segmental information" (Cutler et al. 1997:156).

4.4 Databehandling och dataanalys

Tentamensbetingelserna har alla tre läsåren varit exakt desamma vad gäller testmaterialet, den tid som gått efter terminsstart, yttre betingelser såsom lokal och utrustning etc., urval av försökspersoner samt samma kursplan och lärare. Totala antalet fysiska personer som deltog i vardera tentamen är inte lika stort, vilket styrs av anmälningen till studierna.

De 71 försökspersonerna (18 under läsåret 09-10; 28 för läsåret 10-11 och 25 i läsåret 11-12) hade till uppgift att identifiera och markera vilken ton de uppfattade i 25 kinesiska tvåstaviga ord (50 stavelser). Det bortsågs helt från om markeringen hamnat på fel plats som t.ex. på fel vokal i diftonger/triftonger (5 fall där T1 bör markeras på <u>, alltså korrekt <qiū> och inte <qīu>). Pinyin transkriberades korrekt bortsett från några få fel, t.ex. <ang> istället för <an> (7), och <ong> istället <ung> (8). Sådant räknas inte med i denna studie utan undersökningen befattar sig endast med de felmarkeringar, T' (N = 1035), som försökspersonerna gjorde på de presenterade tonerna ($71 \times 25 \times 2 =$ totalt 3350 responser).

Medeltalet felaktiga identifikationer uppgick till 49,55% för läsåret 09-10, 50,43% för läsåret 10-11 och till 39,40% för läsåret 11-12. Sammanfattningsvis finner vi att den genomsnittliga procentsatsen för felaktigt uppfattade toner under de tre läsåren uppgick till 46,46%, d.v.s. nära nog varannat ord (det rör sig då om fel på antingen den ena eller båda tonerna i ordet). Det tyder på att för läsåret 11-12 har försökspersonernas fel minskat med 10%. En One-way ANOVA där läsårskillarna utgör mellan-subjektfaktor uppdagade inte några signifikanta effekter mellan grupperna [$F(2, 68) = 1.846, p = .166$]. Detta resultat betyder att studenter under de tre årskullarna uppfattar kinesiska toner på samma sätt vilket gör att deltagarna från alla läsåren kan betraktas och därför behandlas som en enda grupp.

Materialet var, som nämndes i avsnittet Stimuli, inte utformat utifrån något statistiskt behov som andra undersökningar. Det som styrde var bara vad den erfarna läraren hade berättat om sina egna upplevelser avseende tonperceptionen och sin erfarenhet.

Variablerna är följande: stavelse (σ_1 och σ_2), tonstimuli (T1, T2, T3, T4 och T0), felresponser (T'1, T'2, T'3, T'4 och T'0), 5 konsonantkategorier, 15 tonkombinationstimuli och slutligen 19 felaktiga tonkombinationsresponser. Variablerna bearbetades statistiskt i IBM SPSS Statistics 22.0.

5 Resultat

Först presenteras en mer detaljerad bild av tonidentifikationerna (5.1) i stavelserna var för sig. Därefter redovisas de vanligaste tonförväxlingarna (5.2) varefter förhållandet mellan konsonantkategorier och tonförväxlingar beskrivs (5.3). Slutligen följer resultat av tonkombinationsförväxlingarna i orden som helhet (5.4).

5.1 Tonidentifikation

5.1.1 Tonidentifikation i σ_1

Tabell 3 Fördelningen av Stimuli (T) $\times$ Felaktiga Responser (T') i σ_1

	T'1		T'2		T'3		T'4		Σ			
	N	%	N	%	N	%	N	%	N	%	M	SD
T1			69	12,43	19	3,42	77	14,05	165	29,91	9,97	5,73
T2	76	13,69			77	14,05	83	14,77	236	42,52	14,17	,54
T3	15	2,52	65	11,71			22	3,96	102	18,20	6,06	4,94
T4	10	1,80	35	6,31	7	1,26			52	9,37	3,12	2,77
Σ	101	18,02	169	30,05	103	18,74	182	32,79	555	100,00		
M		6,00		10,15		6,24		10,93			8,33	
SD		6,67		3,34		6,68		6,04				5,54

Tabell 3, som utgår från bilaga 3, visar det totala antalet felaktiga markeringar (T') som de 71 deltagarna gjort i σ_1 för de 25 stimuli (T) (N = 555, M = 8,33, SD = 5,54).

En One-way ANOVA [$F(3, 8) = 4.234$, $p = .045$] med stimuli som mellan-subjektfaktor visade att skillnaderna var statistiskt signifikanta. Post hoc Tukey HSD visar att T2 (M = 14,17, SD = ,54) får betydligt fler felmarkeringar än T4 (M = 3,12, SD = 2,77) när de utgör stimuli. Detta innebär att T2 är svårast att uppfatta medan T4 är lättast.

En One-way ANOVA med de felaktiga responserna som mellan-subjektfaktor genomförs. Resultatet visar ingen signifikans [$F(3, 8) = .570$, $p = .650$].

5.1.2 Tonidentifikation i σ_2

Tabell 4 Fördelningen av Stimuli (T) $\times$ Felaktiga Responser (T') i σ_2

	T'1		T'2		T'3		T'4		T'0		Σ			
	N	%	N	%	N	%	N	%	N	%	N	%	M	SD
T1			23	4,79	3	,63	24	5,00	5	1,04	55	11,46	2,87	2,35
T2	3	,63			115	23,96	9	1,88	2	,42	129	26,88	6,72	11,50
T3	11	2,29	85	17,71			10	2,08			106	22,08	7,36	8,96
T4	47	9,79	40	8,33	17	3,54			15	3,13	119	24,79	6,20	3,36
T0	3	,63	1	,21			67	13,96			71	14,79	4,93	7,82
Σ	64	13,33	149	31,04	135	28,13	110	22,92	22	4,58	480	100,00		
M		3,34		7,76		9,38		5,73		1,53			5,56	
SD		4,37		7,42		12,71		5,67		1,42				6,76

Tabell 4 som utgår från bilaga 4 visar det totala antalet tonmissuppfattningar i σ_2 som de 71 deltagarna gjort för de 25 stimuli (T) (N = 480, M = 5,56, SD = 6,76).

En One-way ANOVA [$F(4, 13) = .209$, $p = .929$] med stimuli som mellan-subjektfaktor visade ingen signifikant skillnad, vilket innebär att alla stimuli är lika svåra att identifiera. En One-way ANOVA [$F(4, 13) = .662$, $p = .629$] med de felaktiga responserna som mellan-subjektfaktor visade inte heller någon signifikant skillnad.

Tabell 5 Sammanställning av $\sigma 1$ och $\sigma 2$, Independent Sample T-test

stimuli	df	t-test	sig.	$\sigma 1$		$\sigma 2$	
				M	SD	M	SD
T1	5	2,293	,070	9,97	5,73	2,87	2,35
T2	5	1,093	,324	14,17	,54	6,72	11,50
T3	4	-,219	,837	6,06	4,94	7,36	8,96
T4	5	-1,282	,256	3,12	2,77	6,20	3,36
Totala	28	1,179	,248	8,33	5,54	5,56	6,76

Tabell 5 är en sammanställning för både $\sigma 1$ och $\sigma 2$ av de felaktigt uppfattade tonerna. Några signifikanta skillnader föreligger inte mellan $\sigma 1$ ($M = 8,33$, $SD = 5,54$) och $\sigma 2$ ($M = 5,56$, $SD = 6,76$), $t(28) = 1,179$, $p = ,248$ och inte heller mellan olika stimuli. Med avseende på antalet felmarkeringar har alltså de svenska studenterna uppfattat tonerna nästan likadant oavsett stavelse.

5.2 Tonförväxlingar

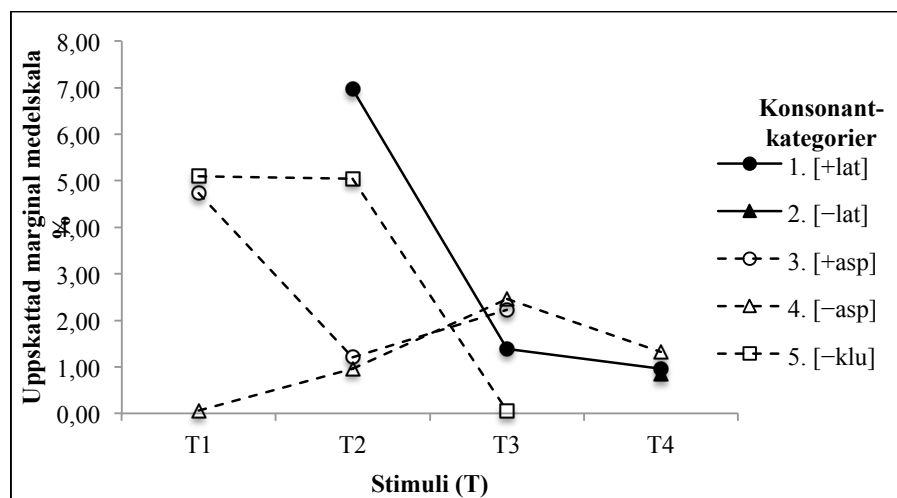
Av tabell 1 framgår det att i $\sigma 1$ är tonförväxlingen $T2 \rightarrow T'4$ (14,77%) den vanligaste, därefter följer $T2 \rightarrow T'3$ och $T1 \rightarrow T'4$ (14,05%), $T2 \rightarrow T'1$ (13,69%), $T1 \rightarrow T'2$ (12,43%) och $T3 \rightarrow T'2$ (11,71%). Genomsnittet för de ovannämnda tonförväxlingarna är signifikant större ($M = 13,45$, $SD = 1,15$) än för det totala antalet ($N = 555$, $M = 8,33$), $t(5) = 28,70$, $p = ,000$.

Tabell 2 visar tydligt att av tonförväxlingarna i $\sigma 2$ är $T2 \rightarrow T'3$ (23,96%) vanligast, därefter följer $T3 \rightarrow T'2$ (17,71%), $T0 \rightarrow T'4$ (13,96%), $T4 \rightarrow T'1$ (9,79%) och $T4 \rightarrow T'2$ (8,33%). De här fem mönstren ($M = 14,75$, $SD = 6,33$) är betydligt större än genomsnittet för det totala antalet ($N = 480$, $M = 5,56$), $t(4) = 5,21$, $p = ,006$.

5.3 Förhållandet mellan konsonantkategorier och tonförväxlingar

5.3.1 Allmänt i $\sigma 1$

Ett upprepat 5×4 ANOVA-test genomfördes där (KON)sonant(K)ategorier ([+lat], [-lat], [+asp], [-asp] och [-klu]) utgör mellan-subjektfaktor och (S)timuli (T1, T2, T3 och T4) inom-subjektfaktor, se figur 9 och bilaga 5. Analysen visar ingen signifikant huvudeffekt för vare sig KONK [$F(4, 10) = 2.508$, $p = .109$] eller för S [$F(2.15, 21.51) = 3.115$, $p = .620$] men däremot föreligger en signifikant interaktion KONK $\times$ S [$F(8.60, 21.51) = 3.240$, $p = .013$].



Figur 9 Förhållandet mellan stimuli och konsonantkategorier i $\sigma 1$

Figur 9 och bilaga 5 visar vertikalt hur de olika konsonantkategorierna (fylld symbol markerar tonande [+son] ljud och ofylld symbol markerar tonlösa [-son]) påverkar olika stimuli (horisonellt).

Tonmissuppfattningarna för stimulus T1 förekommer enbart när initialen är en icke-sonorant; bland dessa påverkar frikativorna [-klu] ($M = 5,10$, $SD = 3,10$) och de aspirerade [+asp] ($M = 4,74$, $SD = 4,76$) i högre grad än icke-aspirerade [-asp] ($M = ,06$, $SD = ,10$). Mot bakgrund av dessa resultat genomfördes ett χ^2 -test mellan KONK ([-klu] och [+asp]) och (T)on(F)örväxlings(M)önstren ($T1 \rightarrow T'2$, $T1 \rightarrow T'3$ och $T1 \rightarrow T'4$). Testet visade att skillnaden mellan TFM och KONK ([+asp] respektive [-klu]) var signifikant, $\chi^2 = (4, N = 165) = 35,624$, $p = ,000$. Detta indikerar att [-klu] oftare är involverad i förväxlingarna $T1 \rightarrow T'2$ och $T1 \rightarrow T'3$ än vad [+asp] är. När det gäller tonförväxlingen $T1 \rightarrow T'4$ påverkar [+asp] kraftigare än [-klu] (se bilaga 5).

Felaktiga identifikationer för T2 sker vanligast när initialen är en sonorant ($M = 6,97$, $SD = 5,15$); därefter följer [-klu] ($M = 5,05$, $SD = 3,44$) medan [+asp] ($M = 1,20$, $SD = 1,61$) och slutligen [-asp] ($M = ,96$, $SD = ,85$) får färre markeringar. Ett χ^2 -test visade att skillnaden mellan KONK (tonande respektive tonlösa) och TFM ($T2 \rightarrow T'1$, $T2 \rightarrow T'3$ och $T2 \rightarrow T'4$) var signifikant, $\chi^2 = (2, N = 236) = 86,168$, $p = ,000$. Detta indikerar att [-son] oftare är involverad i förväxlingen $T2 \rightarrow T'1$ och $T2 \rightarrow T'4$ än vad [+son] är. När det gäller tonförväxlingen $T2 \rightarrow T'3$ påverkar [+son] kraftigare än [-son]. Ett χ^2 -test visade att skillnaden mellan [-son] ([+asp], [-asp] och [-klu]) och TFM var stor, $\chi^2 = (4, N = 120) = 12,571$, $p = ,014$, vilket betyder att [-klu] oftare är involverad i alla TFM ($T2 \rightarrow T'1$, $T2 \rightarrow T'3$ och $T2 \rightarrow T'4$) än vad [+asp] är. Däremot anknyter [+asp] mer till $T2 \rightarrow T'1$ och $T2 \rightarrow T'3$ medan [-asp] mer hör ihop med $T2 \rightarrow T'4$ (se bilaga 5).

När T3 utgör stimulus med olika typer av initialkonsonanter förorsakar de aspirerade ($M = 2,22$, $SD = 3,07$) och icke-aspirerade ($M = 2,46$, $SD = ,85$) fler tonförväxlingar än tonande [+lat] ($M = 1,38$, $SD = ,99$) och [-klu] ($M = ,06$, $SD = ,10$), vilket utgör en distinktiv kontrast till hur KONK förhåller sig till TFM för T2. Ett χ^2 -test visade ingen skillnad mellan KONK (tonande [+son] respektive tonlösa [-son]) och TFM ($T3 \rightarrow T'1$, $T3 \rightarrow T'2$ och $T3 \rightarrow T'4$), $\chi^2 = (2, N = 102) = 3,375$, $p = ,185$. Skillnaden mellan [-son] ([+asp], [-asp] och [-klu]) och TFM var signifikant, $\chi^2 = (4, N = 79) = 14,670$, $p = ,005$. [-klu] är inte involverad med något TFM medan [-asp] är mer associerad med $T3 \rightarrow T'1$ och $T3 \rightarrow T'4$. Slutligen är [+asp] mer förknippad med $T3 \rightarrow T'2$ (se bilaga 5).

När det gäller T4 påverkar [-asp] ($M = 1,32$, $SD = 1,10$) tonperception mer än [+lat] ($M = ,96$, $SD = ,45$) och [-lat] ($M = ,84$, $SD = 1,30$). Ett χ^2 -test visade inte någon skillnad mellan KONK (tonande [+son] respektive tonlösa [-son]) och TFM ($T4 \rightarrow T'1$, $T4 \rightarrow T'2$ och $T4 \rightarrow T'3$), $\chi^2 = (2, N = 52) = 1,900$, $p = ,387$. Däremot var skillnaden mellan [+lat] och TFM stor, $\chi^2 = (2, N = 30) = 7,089$, $p = ,029$, vilket innebär att [+lat] förknippas mer med $T4 \rightarrow T'1$ och $T4 \rightarrow T'3$ än [-lat] medan å andra sidan [-lat] drabbar $T4 \rightarrow T'2$ i större utsträckning (se bilaga 5).

5.3.2 Allmänt i $\sigma 2$

Ett upprepat 5×5 ANOVA-test, där KONK ([+lat], [-lat], [+asp], [-asp] och [-klu]) utgör mellan-subjektfaktor och med S ($T1$, $T2$, $T3$, $T4$ och $T0$) som inom-subjektfaktor visade ingen signifikant huvudeffekt för vare sig KONK [$F(4, 15) = 1.201$, $p = .351$] eller för S [$F(2.41, 36.09) = .524$, $p = .629$] och heller ingen signifikant interaktion KONK $\times$ S [$F(9.62, 36.09) = .936$, $p = .511$], se bilaga 6.

5.4 Tonkombinationer

Stimuli utgörs av 15 olika tonkombinationer (TKS) som kan utgöras av vilken som helst av 19 felaktiga tonkombinationsresponser (FTKR), se bilaga 7. Ingen signifikant skillnad föreligger mellan de olika TKS genom en one-way ANOVA, [$F(14, 131) = .773$, $p = .696$] men

däremot mellan de olika FTKR, [$F(18, 127) = 1.698, p = .048$].

Tabell 6 Tukey HSD^{a,b} av FTKR

FTKR	N	Subset for alpha = 0.05	
		1	2
1+0	7	,20700	
4+0	4	,21725	
2+0	6	,35433	
1+1	7	,38643	
3+1	8	,44675	
3+2	9	,47222	
1+4	9	,53678	
2+2	9	,55822	
4+4	11	,56227	
2+1	9	,57956	
3+0	4	,65225	
2+3	7	,71786	,71786
4+2	12	,74883	,74883
3+4	7	,81443	,81443
4+3	8	,82112	,82112
4+1	6	,85367	,85367
2+4	13	,89185	,89185
1+2	5	1,33340	1,33340
1+3	5		2,29940
Sig.		,560	,063

Means for groups in homogeneous subsets are displayed.

a. Uses Harmonic Mean Sample Size = 6,908.

b. The group sizes are unequal. The harmonic mean of the group sizes is used. Type I error levels are not guaranteed.

Detta resultat illustrerar att försökspersonerna, oberoende av vilka tonkombinationer som utgör stimuli, tycks uppfatta två kategorier där mönstren från de svenska ordaccenterna assimilerats till de kinesiska tonkombinationerna.

6 Slutsats och Allmän diskussion

6.1 Slutsats

Trots att det inte föreligger någon signifikant skillnad mellan σ_1 och σ_2 angående det totala antalet felmarkeringar skiljer sig tonuppfattningsresultaten för σ_1 i de 25 tvåstaviga testorden från σ_2 . I σ_1 är det T2 som de svensktalande studenterna har svårast att identifiera medan T4 är lättast. I σ_2 bereder alla stimuli försökspersonerna svårigheter i lika hög grad.

Tonförväxlingar sker på olika sätt mellan σ_1 och σ_2 . Problematiska är – och tycks förbli – förväxlingarna T2→T'3 och T3→T'2 men för de svensktalande försökspersonerna är dessa mönster mindre dominerande än de verkar vara för personer med andra modersmål.

Konsonantkategorierna inverkar på tonförväxlingarna i σ_1 på så sätt att olika konsonanter kan knytas till vissa tonförväxlingar. Inga sådana samband kan konstateras för σ_2 .

Det verkar vara så att våra svensktalande studenter, oberoende av vilken kinesisk tonkombination som utgör stimulus, uppfattar att tonmönster låter som de svenska ordaccenterna.

6.2 Allmän diskussion

Forskning om hur svenskar uppfattar kinesiska toner är sällsynt. Det är svårt att jämföra med tidigare forskningsresultat som grundar sig på personer som har annat L1 än svenska, det må vara kinesiska eller annat språk. För att på ett relativt objektivt sätt observera hur svenska vuxna nybörjarstudenter uppfattar kinesiska toner insamlades så mycket nya data som möjligt. Denna metod ställs mot de flesta traditionella perceptionsexperiment där forskarna utgår från

en hypotes och därefter utformar testet som slutligen genomförs. De resultat som har presenterats här får antas gälla för svenskar som har den västsvenska varianten av svenska som modersmål.

Testorden i materialet var inte tillrättalagda för den nu genomförda studien utan valda för att utgöra en väl lämpad tentamen. Valet av orden och den ljuddämpade (men inte helt tysta) inspelningslokalen kan motiveras med att i verkligheten hör man knappast rent syntetiserade och isolerade ljud/enskilda stavelser i en totalt tyst miljö, detta gäller speciellt för dem som studerar språk. Dessutom tillåter undervisningens ekonomi inte att ta ett helt pass för att lyssna på yttranden på det här sättet. En undersökning av den dynamiska tonperceptionsprocessen kräver att materialet utgörs av tvåstaviga ord. I det ovanstående kontrasterar den aktuella studien mot de flesta andra experiment som dessutom oftast koncentrerats till F0 i enstaviga ord.

Syftet med denna uppsats var att ta reda på hur svenskar uppfattar kinesiska tvåstaviga ord. Det faktum att nästan alla transkriberat rätt kan dels tolkas som att lyssnarna omedvetet koncentrerar sig mest på segmenten och, sett från andra hållet, dels tyda på att ljuden påverkar tonperceptionen. Det ovan anförda kan tolkas som att det skulle vara bättre för nybörjarna att först möta segment, ordprosodi och intonation istället för att direkt, stackato, träna tonernas förlopp med handrörelser. Detta ligger också i linje med Kiriloffs (1969) tidiga experiment där det visade sig att tonidentifieringarna blev bättre när försökspersonerna inte behövde transkribera.

Trots att materialet är relativt litet och metoden skiljer sig från tidigare forskning har vi ändå funnit likheter och skillnader som att T2 är den svåraste tonen att uppfatta medan T4 är lättast när det gäller första stavelsen men att denna skillnad inte återfinns i den andra stavelsen. Om vi följer tidigare resonemang skulle man kunna säga att T4 verkar uppfattas som vara påståendesats (So & Best 2014), särskilt i slutet av ett yttrande (Broselow et al. 1987). Någon förklaring till att T2 är svårast att uppfatta finner man sällan i litteraturen. Dessutom sker den kinesiska stigande rörelsen inom ett väldigt kort tidsintervall – kan svenskars ovana vid sådant kunna ligga bakom svårigheterna? Enligt transferteorin borde svenska studerande å andra sidan inte ha detta problem eftersom tonförloppet i den svenska akuta accentens uppträdande i $\sigma 1$ liknar den kinesiska T2-kurvan. För svensk grav accent ser tonförloppet i bägge stavelserna ut som en kinesisk T4 (jfr. figurerna 1 och 7). Eller beror detta på att T2 och T4 är mer utsatta för påverkan från L1 eftersom de inte är lika prosodiskt markerade (Abrahamsson 2004:89-92)? Varför är T4 lättare än T2 trots att båda tonernas egenskaper finns i svenskan? Uppmärksammar försökspersonerna toppigheterna inom onset respektive offset mer än tonhöjdsskillnaden över tid och därmed förlorar känslan för tonförloppens riktningar i deras helhet? Eller beror också detta på att nybörjarna inte är medvetna om att/hur deras eget språk påverkar det främmande språket? Varför föreligger en skillnad mellan $\sigma 1$ och $\sigma 2$?

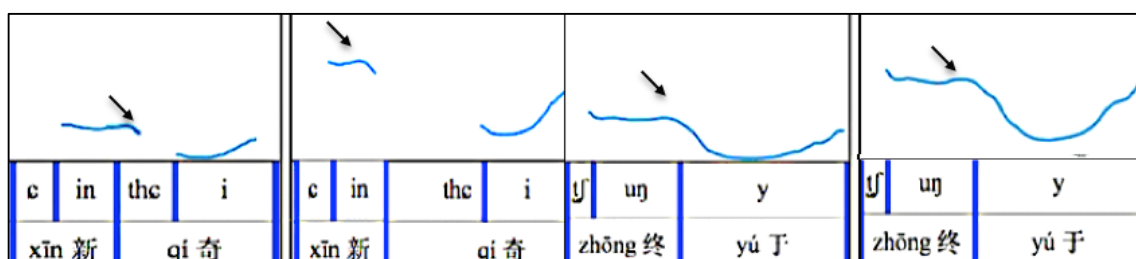
Vi har också sett att våra försökspersoner, oavsett tonstimulus, uppfattar tonerna i $\sigma 2$ likadant vilket innebär att det är svårare att uppfatta toner i $\sigma 2$ än $\sigma 1$. Detta stämmer inte med vad Lin & Wang (2007a, se ovan 2.2.2) definierar som *auditory illusion* (och alltså $\sigma 1$ borde felmarkeras oftare, s. 15). Som stimulus i tvåstaviga ord använde deras undersökning enbart T1. Vid en begränsning av vårt material till de stimuli där T1 står i $\sigma 1$ verifierar vårt resultat deras fynd: T1 i $\sigma 1$ ($M = 9,97$, $SD = 5,73$), se tabell 3, felmarkeras betydligt oftare än T1 i $\sigma 2$ ($M = 2,87$, $SD = 2,35$), se även tabell 4.

I våra resultat förekommer här flera förväxlingsmönster gällande bägge stavelserna jämfört med tidigare forskningsresultat. Det förekommer ingen signifikant skillnad förväxlingsmönstren emellan. Men T2→T'3 är ett problem också för svenskarna. Om vi följer Shen & Lin (1991) verkar våra resultat stämma överens med deras vändpunkt (*turning point*). Som exempel kan tas ord nr.12-劳动-<láodòng> 'arbeta' där $\sigma 1$ består av 劳 <láo> som inleds med sonoranten [l] och bär en T2: här sker hos nästan alla en förväxling till T'3 och således missuppfattas stimulus som ordet 老动 <lǎodòng> 'kan inte vara stilla'. Mot detta kontrasterar det

aktuella materialet med att i de ord med T2 i σ_1 som inleds med tonlös initial förväxlas den inte till T'3 utan till antingen T'1 eller T'4. Härav kan vi sluta oss till att enbart vändpunkten inte helt förklarar förväxlingen T2→T'3.

Detta står inte i konflikt till Xu & Xu (2003) som menar att aspiration ger upphov till en lägre F0. De fann att aspiration påverkar F0 beroende på den ton som utgör stimulus, särskilt då för skillnaden mellan T2 och T3. Detta stämmer inte med det aktuella resultatet för stavelsen 体 <ti> med en T3 i ord nr.16-体验-<tiyàn>-‘erfaren’; initialen är aspirerad och förväxlas till T'2. Intressant är att försökspersonerna, trots aspirationen, ändå uppfattar T3 (som är L) som en hög ton (H). Hombert (1978) verkar här vara närmare en sanning när han påpekar att en tonlös initial konsonant resulterar i att tonen uppfattas som en H. För övriga förväxlingar (T1→T'4, T1→T'2, T2→T'1) skulle vi också kunna säga att olika konsonantkategorier spelar roll gällande tonförväxlingar.

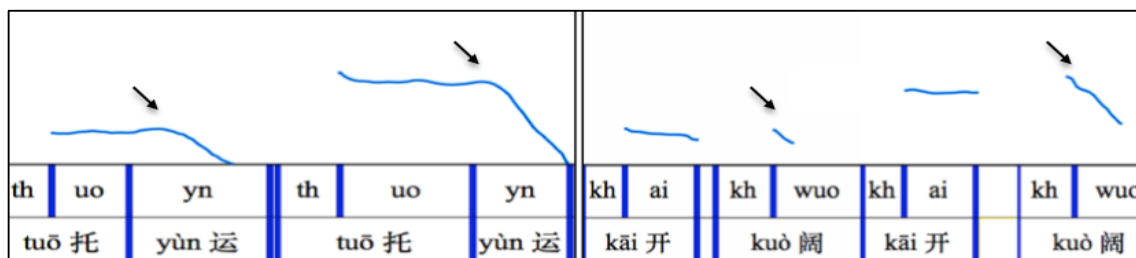
I σ_2 sker tonförväxlingar (T2→T'3 och T3→T'2 är de vanligaste även här) men dessa förväxlingar kan inte knytas till någon konsonantkategori. Detta beror antagligen på att den föregående tonen (i σ_1) påverkar den efterföljande.



Figur 10 F0-kontur (överst), IPA (mellan) av testorden nr. 13-新奇-<xīnqí>-[ɕin˥-tɕʰi˥]-‘nyfiken’; manlig talare respektive kvinnlig talare och nr. 17-终于-<zhōngyú>-[ʈʂʊŋ˥-y˨˥]-‘äntligen’; manlig talare (till vänster) respektive kvinnlig talare (till höger).

Figur 10 visar ord nr. 13-新奇-<xīnqí>-[ɕin˥-tɕʰi˥]-‘nyfiken’ (man till vänster och kvinna till höger) och nr. 17-终于-<zhōngyú>-[ʈʂʊŋ˥-y˨˥]-‘äntligen’. Den andra stavelsen i ord 13 inleds med en (tonlös) frikativa medan den andra stavelsen i ord 17 inleds av en halvvokal (tonande). Bägge orden har som stimulus i andra stavelsen en T2 som i båda fallen missuppfattas och förväxlas till T'3. T1 i den första stavelsen (gäller bägge orden) sjunker för att komma ner till den kommande låga tonhöjd som utmärker början av T2. Förväxlingen kan tänkas bero på vändpunkten i tonkurvan.

Förväxlingarna T4→T'1 och T4→T'2 i σ_2 illustreras i figur 11 med orden nr. 23-托运-<tuōyùn>-[tʰuo˥-yn˥]-‘frakt’ och nr. 7-开阔-<kāikuò>-[kʰai˥-kʰwo˥]-‘vidsträckt’. Ord 23 inleds av en vokal utan konsonant medan den andra stavelsen i ord 7 inleds med en aspirerad klusil (tonlös).



Figur 11 F0-kontur (blå heldragen linje) av testorden nr. 23-托运-<tuōyùn>-[tʰuo˥-yn˥]-‘frakt’; manlig talare respektive kvinnlig talare och nr. 7-开阔-<kāikuò>-[kʰai˥-kʰwo˥]-‘vidsträckt’; manlig talare respektive kvinnlig talare.

I ord nr. 23 (till vänster) startar F0 med ett första maximum där första stavelsens triftong [ʷuo] börjar respektive [ai] i ord nr. 7 (till höger). Bägge orden visar samma konfiguration som början av en svensk grav accent. F0 sjunker däremot inte (som den gör i svenskan) i ord 23 utan fortsätter behålla sin höjd ända till den andra stavelsen [yn] p.g.a. att *onset* för både T1 och T4 ligger lika högt. Det kan tolkas som att våra lyssnare uppfattar en högsta tonhöjd och sedan inte lyssnar vidare (annars skulle de inte gjort misstagen T4→T'1); de "har fått tillräcklig information", nämligen tontoppigheten i den andra stavelsen.

När det gäller ord nr. 7 sjunker F0 en aning i slutet av T1 i σ1, förmodligen som förberedelse för den efterföljande aspirationen. Avslutningen av första stavelsens T1 stiger till den punkt där T4 i den andra stavelsen, [ʷuo], börjar. Även detta skulle kunna tolkas som att våra lyssnare uppfattar en högsta tonhöjd och sedan inte lyssnar vidare (annars skulle de inte gjort misstagen T4→T'2); återigen har de "fått tillräcklig information", nämligen tontoppigheten i den andra stavelsen. Beteendet att bara uppmärksamma den högsta tonhöjden liknar det som engelsmännen och nederländarna visade i Leathers studie (1987). Mot detta ska ställas att när det gäller kinesisk tonperception är det hela rörelsen som spelar roll (Gandour 1984).

Testordens tonkombinationer i tvåstaviga ord som helhet spelar inte någon roll. De felaktiga responserna på tonkombinationerna kan delas upp i två kategorier, vilket, som påpekats tidigare, kan tolkas som att de svenska tonaccenterna styr hur kinesiskan uppfattats. Felaktiga responser på tonkombinationer kan tydas som att orden uppfattats fel p.g.a. inverkan från svenska accenter. De kinesiska F0-kurvorna uppvisar inte alltid samma bild som de svenska tonaccenterna, se figur 11. Det kan inträffa att en toppighet i ett kinesiskt tvåstavigt ord hamnar på andra stavelsen, vilket kan bero på segmentella strukturer vid stavelsegränserna. Då är frågan om vilken stavelse som ska anses vara mer prominent än den andra. Här behövs mer akustisk analys.

7 Slutord och framtida forskning

I denna uppsats rapporteras hur svenska vuxenstuderande på nybörjarnivå uppfattar tonerna i kinesiska tvåstaviga ord. Vi har visat att L2-tonperceptionen inte längre kan baseras på F0 och dess akustiska parametrar i enstaviga ord. Inte heller kan försökspersonernas missuppfattningar av tonerna skyllas på enkel ömsesidig assimilation mellan kinesiska enstaviga toner och svenskans intonation. Vi finner oss stå inför en komplex och komplicerad dynamisk process. Denna börjar redan med att olika konsonantkategorier inverkar på F0 eftersom det kinesiska ljudsystemet skiljer sig från det svenska. När sedan tonförloppet i σ1 övergår till σ2 påverkas den efterföljande tonen (tonsandhi). En möjlighet föreligger att svenska tonaccenter i viss mån assimileras till den kinesiska tonkombinationen p.g.a. att det som utgör skillnaderna mellan accenterna är tontoppigheten i de svenska orden.

Trots att materialet är begränsat återkommer de felaktiga responsmönstren årskull efter årskull. Därför kan resultaten inte ignoreras, de är rentav betydelsefulla både för L2-kinesisk undervisning och vidare jämförelse av ordprosodin i svenskan och kinesiskan.

Om de olika konsonantkategoriernas inverkan på tonperceptionen kan säkerställas behöver undervisningen inte längre i första hand lägga tonvikten på träning av tonerna. Istället gäller det att få studenterna att perceptoriskt skilja ett ljud från ett annat. Tonträningen skulle inte längre fokuseras på enstaviga ord ledsagade av gester eller på diakritisk beteckning i kurslitteratur (Ahlgren & Löfstedt 1973; Björkstén & Erlandsson 2007; Liu 2009; Malmqvist 1979; T'Ung & Pollard 1983). Däremot bör energin läggas på tonala, temporala och dynamiska skillnader mellan tonerna som t.ex. att T2 i stavelserna <dí/tí> är ljusare än i T3. Vad beträffar tonkombinationer bör man ta upp minimala par (eller rentav trillingar) som skiljs från varandra enbart med ton t.ex. <dárén> 'intellegent', <dären> 'slå någon', <dårén> 'ämbetsmän'. Dessa uppvisar likheter med de skilda prosodierna i svenskans <Milan> (accent I) och <milan> (accent II). Rörande betoning skulle man kunna ta upp <daren> igen och ställa frågor t.ex. vilken stavelse som uppfattas som prominent, lite i likhet med svenska partikelverb som

<`hälsa på> ‘säga hej’ och <hälsa `på> ‘besöka’. Sådana ordpar kan också tränas i kontext för att jämföra ordprosodin i sig med hur prosodemen förändras vid stavelsegränserna.

Förutom praktisk undervisning behövs framtida forskning. Dels bör konsonanternas inverkan på tonuppfattningen i enstaviga ord undersökas vidare – varför inträffar sådant enbart i första stavelsen och hur påverkas språk med andra klusilsystem än det svenska? Akustiska jämförelser mellan kinesiska och den studerandes tidigare inlärd språk måste genomföras för att kunna lotsa den lärande fram. Syftet med en framtida forskning är inte att bara bygga teoretiska modeller utan att utgå från kontrastiv lingvistik, pedagogisk verksamhet och erfarenhet, datainsamling och didaktiska försök. Det skulle vara intressant att få reda på hur kineser uppfattar och tolkar de svenska accenterna. Mennen (2015:172) påstår att “[T]o date, no model has been proposed that exclusively deals with and makes predictions of the relative difficulty of producing and perceiving non-native intonation, although some recent attempts have been made to extend the PAM-L2 to the perception of lexical tones (So and Best, [2010, 2011], 2014)”.

Referenser

- Abrahamsson, N. (2004). Fonologiska aspekter på andraspråksinläring och svenska som andraspråk In K. Hyltenstam & I. Lindberg (eds.) *Svenska som andraspråk : i forskning, undervisning och samhälle*. Lund: Studentlitteratur, 97-116.
- Abrahamsson, Niclas (2009). *Andraspråksinläring*. Lund: Studentlitteratur.
- Ahlgren, Inger & Jan-Ingvar Löfstedt (1973). *Lärobok i folkkinesiska*. Stockholm: Skriptor.
- Bannert, Robert (1994). *På väg mot svenskt uttal*. Lund: Studentlitteratur.
- Best, Catherine T (1995). A direct realist view of cross-language speech perception In W. Strange (ed.) *Speech perception and linguistic experience : issues in cross-language research*. Baltimore: York Press, 171-206.
- Bidelman, Gavin M, Jackson T Gandour & Ananthanarayan Krishnan (2011). Cross-domain effects of music and language experience on the representation of pitch in the human auditory brainstem. *Journal of Cognitive Neuroscience* **23**, 425-434.
- Björkstén, Johan & Anna Erlandsson (2007). *Kinesiska : språket i Mittens rike*. Stockholm: Bilda.
- Broselow, Ellen, Richard R Hurtig & Catherine Ringen (1987). The perception of second language prosody. *Interlanguage phonology: The acquisition of a second language sound system*, 350-361.
- Bruce, Gösta (2005). Intonational prominence in varieties of Swedish revisited. In S.-A. Jun (ed.) *Prosodic typology : the phonology of intonation and phrasing*. Oxford: Oxford University Press, 411-428.
- Bruce, Gösta (2012). *Allmän och svensk prosodi*. Lund: Studentlitteratur.
- Catford, John Cunnison (2001). *A practical introduction to phonetics*. Oxford: Oxford University Press.
- Chao, Yuen Ren (1968). *A grammar of spoken Chinese*. Berkeley.
- Chen, Ao (2013a). Universal biases in the perception of Mandarin tones: From infancy to adulthood. LOT dissertation series **315**.
- Chen, C. M (2013b). Gestures as tone markers in multilingual communication. . *Research in Chinese as a second language*. 143.
- Chuang, C.K., S. Hiki, T. Sone & T. and Nimura (1972). The Acoustical features and perceptual cues of the four tones of standard colloquial Chinese Proceedings of the 7th International Congress of Acoustics **3**, 297-300.
- Clark, John Ellery, Colin Yallop & Janet Fletcher (2007). *An introduction to phonetics and phonology*. Oxford ; Malden, Ma.: Blackwell.
- Cutler, Anne, Delphine Dahan & Wilma Van Donselaar (1997). Prosody in the comprehension of spoken language: A literature review. *Language and speech* **40**, 141-201.
- Ding, Hongwei (2012). Perception and production of Mandarin Disyllabic Tones by German Learners. Tongji University Press, Shanghai, 378-381.
- Ding, Hongwei, Oliver Jokisch & Rüdiger Hoffmann (2010). Perception and Production of Mandarin Tones by German Speakers. *Speech Prosody*, Chicago.
- Duanmu, San (2007). *The phonology of standard Chinese*. Oxford: Oxford Univ. Press.
- Engstrand, Olle (2004). *Fonetikens grunder*. Lund: Studentlitteratur.
- Flege, J. E. (1995). Second language speech learning: theory, findings, and problems. In W. Strange (ed.) *Speech perception and linguistic experience : issues in cross-language research*. Baltimore: York Press, 233-277.
- Flege, James Emil (1987). The production of "new" and "similar" phones in a foreign language: Evidence for the effect of equivalence classification. *Journal of phonetics* **15**, 47-65.
- Gandour, J. (1983). Tone perception in Far Eastern languages *Journal of Phonetics* **11**, 149-175.
- Gandour, J. (1984). Tone dissimilarity judgments by Chinese listeners *Journal of Chinese Linguistics* **12-2**, 235-261.
- Garlén, Claes (1988). *Svenskans fonologi : [i kontrastiv och typologisk belysning]*. Lund: Studentlitteratur.
- Gass, Susan M. & Larry Selinker (2008). *Second language acquisition : an introductory course*. New York: Routledge/Taylor and Francis Group.
- Gottfried, Terry L & Traci L Suiter (1997). Effect of linguistic experience on the identification of Mandarin Chinese vowels and tones. *Journal of Phonetics* **25**, 207-231.
- Guo, Jinfu (1993). Survey of Chinese tone and intonation. Beijing: Beijing Language University Press. 郭锦桴 《汉语声调语调纲要与探索》，北京语言学院出版社，北京。 <Hànyǔ shēngdiào yǔdiào chānyào yǔ tànsoù, Běijīng Yǔyánxuéyuàn chūbǎnshè, Běijīng>.
- Gårding, Eva (1998). Intonation in Swedish. In A. Di Cristo, D. Hirst & E. Gårding (eds.) *Intonation systems : a survey of twenty languages*. Cambridge: Cambridge Univ. Press, 112-130.
- Gårding, Eva, Paul Kratochvil, Jan-Olof Svantesson & Jialu Zhang (1986). Tone 4 and Tone 3 Discrimination in Modern Standard Chinese. *Language and speech* **29**, 281-293.
- Hallé, Pierre A., Yueh-Chin Chang & Catherine T. Best (2004). Identification and discrimination of Mandarin Chinese tones by Mandarin Chinese vs. French listeners *Journal of Phonetics* **32**, 395-421.
- Helgason, Pétur & Catherine Ringen (2008). Voicing and aspiration in Swedish stops. *Journal of Phonetics* **36**, 607-628.
- Hombert, Jean-Marie (1978). Consonant Types, Vowel Quality and Tone. In V. A. Fromkin (ed.), 77-111.
- House, David (1990). *Tonal perception in speech*. Diss Lund Univ, Lund Univ. Press ;

- Chartwell-Bratt.
- Howie, John Marshall (1976). *Acoustical studies of Mandarin vowels and tones*. Cambridge: Cambridge U.P.
- Hu, Guohua (2012). Chinese and Swedish Stops in Contrast. Gothenburg Department of Philosophy, Linguistics and Theory of Science, 77-80.
- Karlgren, Bernhard (1965). *Kina i tal och skrift : [en introduktion till kinesiska språket]*. Stockholm,.
- Kirilloff, C. (1969). On the auditory perception of tones in Mandarin *Phonetica* **20**, 63-67.
- Klatt, D. (1973). Discrimination of fundamental frequency contours in synthetic speech duplicaitons for models of pitch perception *Journal of the Acoustical Society of America* **53**, 8-16.
- Kratochvil, Paul (1998). Intonation in Beijing Chinese. In A. Di Cristo, D. Hirst & E. Gårding (eds.) *Intonation systems : a survey of twenty languages*. Cambridge: Cambridge Univ. Press, 421-435.
- Leather, J. (1987). F0 pattern inference in the perceptual acquistition of second language tone In A. J. J. Leather (ed.), 59-80.
- Lee, Chao-Yang, Liang Tao & Z. S. Bond (2010). Identification of Acoustically Modified Mandarin Tones by Non-native Listeners. *Language and speech* **53**, 217-243.
- Lee, Wai-Sum & Eric Zee (2003). Standard Chinese (Beijing). *Journal of the International Phonetic Association* **33**, 109-112.
- Lee, Wai-Sum & Eric Zee (2008). Prosodic characteristics of the neutral tone in Beijing Mandarin. *journal of Chinese Linguistics*.
- Li, Charles N. & Sandra A. Thompson (1976). The acquisition of tone in Mandarin -speaking children. *Journal of Child Language* **4**, 185-199.
- Liang, Zhian (1963). Findings of Chinese tone perception. *Journal of physiology* **26**. pp. 85-91. 梁之安 《汉语普通话中声调的听觉辨认依据》，《生理学报》第 26 期，第 85-91 页。 <Hanyǔ pǔtōnghuà-zhōng shēngdiào de tīngjué biànrèn yījù, Shēnglǐ xuébào, dì 26 qī, dì 85-91 yè.>
- Lightbown, Patsy M. & Nina Spada (2006). *How languages are learned*. Oxford: Oxford Univ. Press.
- Lin, Tao (1985). Test and preliminary discussion on T0 in Beijing spoken Chinese. In Lin Tao's linguistic works, pp. 120-141, Beijing: The Commercial Press. 林焘《探讨北京话轻声性质的初步实验》，选自《林焘语言学论文集》，第 120-141 页，商务印书馆，北京。 <Tàntǎo Běijīnghuà qīngshēng xìngzhì de chūbù shíyàn, xuǎnzi Lín Táo yǔyánxué lùnwénjí, dì 120-141 yè, Shāngwù yīnshūguǎn, Běijīng.>
- Lin, Tao (2007a). 林焘《声调感知问题》，选自《林焘语言学论文集》，第 142-155 页，商务印书馆，北京。 <Shēngdiào gǎnzhī wèntí, xuǎnzi Lín Táo yǔyánxué lùnwénjí, dì 142-155 yè, Shāngwù yīnshūguǎn, Běijīng.>
- Lin, Yen-Hwei (2007b). *The Sounds of Chinese*. Cambridge University Press.
- Liu, Jiangna (2006a). The study on the analysis of Chinese error tones by Korean student on intermedia level. M. A. Dissertation. Tianjin Normal University. 刘江娜《中级汉语水平的韩国留学生声调偏误问题研究》，硕士论文，天津师范大学。 <Zhōngjí hànyǔ shuǐpíng de hánguó liúxuéshēng shēngdiào piānwù wèntí yánjiū, Shuòshì lùnwén, Tiānjīn shīfàn dàxué.>
- Liu, Lih-Cherng, Wu-Ji Yang, Hsiao-Chuan Wang & Yueh-Chin Chang (1989). Tone recognition of polysyllabic words in Mandarin speech. *Computer Speech & Language* **3**, 253-264.
- Liu, Xiaojun (2006b). The Chinese tone error analysis by Vietnam students. M. A. Dissertation. Guangxi University. 刘晓军《越南留学生汉语声调偏误实验分析》，硕士论文，广西大学。 <Yuè'nán liúxuéshēng hànyǔ shēngdiào piānwù shíyàn fēnxī, Shuòshì lùnwén, Guǎngxī dàxué.>
- Liu, Yuehua (2009). *Integrated Chinese = [Zhong wen ting shuo du xie]. Level 1, part 1*. Boston: Cheng & Tsui Company.
- Maddieson, Ian & Sandra Ferrari Disner (1984). *Patterns of sounds*. Cambridge: Cambridge U.P.
- Malmqvist, Göran (1979). *Nykinesisk fonetik*. Stockholm: Inst. för orientaliska språk, Stockholms univ.
- Manne, Ineke (2015). Prosody and Language in Contact L2 Acquisition, Attrition and Languages in Multilingual Situations. In E. Delais-Roussarie, M. Avanzi & S. Herment (eds.): SpringerLink (Online service), 171-188.
- McAllister, Robert (1996). Aspects of the relation between the production and perception of a second language *TMH-QPSR* **37**, 69-72.
- McAllister, Robert (1998). *Talkommunikation*. Lund: Studentlitteratur.
- Sebastián - Gallés, Núria (2005). Cross - Language Speech Perception. *The handbook of speech perception*, 546-566.
- Shen, Beilei (2006). The study on word stress patterns of disyllable non-neutral words in Mandarin. M. A. Dissertation. East China Normal University. 沈倍蕾《普通话双音节非轻声词的轻重格式研究》，硕士论文，华东师范大学。 <Pǔtōnghuà shuāngyīnjié fēiqīngshēngcí de qīngzhòng géshì yánjiū, Shuòshì lùnwén, Huádōng shīfàn dàxué.>
- Shen, Sussan Xiaonan (1989). Toward a register approach in teaching Mandarin tones. *Journal of Chinese Language Teachers Association* **24**, 27-47.
- Shen, Xiaonan Susan & Maocan Lin (1991). A Perceptual Study of mandarin Tones 2 and 3. *Language and speech* **1991, 34**, 145-156.
- Sjögren, Peter A. (1979). *Termer i allmän språkvetenskap : ett systematiskt lexikon*. Stockholm: Akademilitt.

- So, Connie K & Catherine T Best (2014). Phonetic influences on English and French listeners' assimilation of Mandarin tones to native prosodic categories. *Studies in Second Language Acquisition* **36**, 195-221.
- Strange, Winifred (1995). Cross-language studies of speech perception In W. Strange (ed.) *Speech perception and linguistic experience : issues in cross-language research*. Baltimore: York Press, 3-45.
- T'Ung, Ping-cheng & David E. Pollard (1983). *Colloquial Chinese. Character text for Colloquial Chinese : simplified character version*. London: Routledge & Kegan Paul.
- Trask, Robert Lawrence (1995). *A Dictionary Of Phonetics And Phonology*.
- Wang, Lijia & Tao Lin (1992). *Course of phonetics*. Beijing University Press, Beijing. 王理嘉、林焘 《语音学教程》，北京大学出版社。<Yǔyīnxué jiàochéng, Běijīng dàxué chūbǎnshè, Běijīng.>
- Wang, Lijia, Jianming Lu, Huaiqing Fu, Zhen Ma & Peicheng Su (2004). *Modern Chinese* 王理嘉、陆剑明、符淮清、马真、苏培成 《现代汉语》，商务印书馆，北京。<Xiàndài hànyǔ, Běijīng, Shāngwù yīnshūguǎn>.
- Wang, Youmin (1998). Analysis of tone markings of disyllabic words by Hungarian students. *Journal of Chinese Teaching in the World* **2**, pp. 102-106. 王又民 《匈牙利学生汉语双音词声调标注量化分析》，《世界汉语教学》，第 2 期，第 102-106 页。<Xiōngyáli xuéshēng hànyǔ shuāngyīncí shēngdiào biāozhù liàngguà fēnxī, Shìjiè hànyǔ jiàoxué, dì 2 qī, dì 102-106 yè.>
- Wang, Yue, Michelle M. Spence, Allard Jongman & Joan A. Sereno (1999). Training American Listeners to Perceive Mandarin Tones. *Journal of the Acoustical Society of America* **106**, 3649-3658.
- Wang, Yunjia (2004). The effects of pitch and duration on the perception of the neutral tone in standard Chinese, *Journal of acoustics* **29**, pp. 453-461. 王韞佳 《音高和时长在普通话轻声知觉中的作用》，《声学学报》第 29 期，第 453-461 页。<Yīngāo hé shícháng zài pǔtōnghuà qīngshēng-zhōngde zuòyòng, dì 29 qī, dì 453-461 yè.>
- Whalen, Douglas H & Yi Xu (1992). Information for Mandarin tones in the amplitude contour and in brief segments. *Phonetica* **49**, 25-47.
- White, Carolyn M (1981). Tonal perception errors and interference from English intonation. *Journal of the Chinese Language Teachers Association* **16**, 27-56.
- Wu, Zongji & Maocan Lin (1989). *Overview of Experimental Phonetics*, Higher Education Press, Beijing. 吴宗济、林茂灿 《实验语音学概要》 高等教育出版社，北京。<Shíyàn yǔyīnxué gǎiyào, Gāoděng jiàoyù chūbǎnshè, Běijīng.>
- Xu, Ching X & Yi Xu (2003). Effects of consonant aspiration on Mandarin tones. *Journal of the International Phonetic Association* **33**, 165-181.
- Xu, Yi (1997). Contextual tonal variations in Mandarin. *Journal of Phonetics* **25**, 61-83.
- Yip, Moira Jean Winsland (2002). *Tone*. Cambridge: Cambridge University Press.
- Zhang, Lijuan (2012). *The perception and pronounce of Chinese Mandarin - Tones by English and German native speakers*. M. A. Dissertation. Beijing University. 张丽娟 《英德母语者对普通话声调的感知和发音》，硕士论文，北京大学。<Yīngdé mǔyǔzhě duì pǔtōnghuàshēngdiào de gǎnzhī hé fāyīn, Shuòshì Lùnwén, Běijīng dàxué.>

Bilaga 1

De 25 tvåstaviga orden

Tvåstaviga ord		TK	Pinyin	IPA	Svenska
σ1	σ2				
1 黄	1 河	22	huánghé	xwan ^h -xɿ ^h	Gula floden
2 凉	2 爽	23	liángshuǎng	ljan ^h -ɕwan ^h	kallt, svalt, uppfriskande
3 秋	3 风	11	qiūfēng	tɕ ^h iou ^h -fəŋ ^h	höstvind
4 旅	4 客	34	lǚkè	ly ^h -k ^h ɿ ^h	resenär
5 广	5 播	31	guǎngbō	kwan ^h -pwuo ^h	sändning
6 立	6 刻	44	likè	li ^h -k ^h ɿ ^h	per omgående, genast
7 开	7 阔	14	kāikuò	k ^h ai ^h -k ^h wuo ^h	vidsträckt
8 吸	8 引	13	xīyǐn	ɕi ^h -jin ^h	lockande, attrahera
9 照	9 相	44	zhàoxiàng	tʃau ^h -ɕaŋ ^h	Fotografera
10 寻	10 找	23	xúnzhǎo	ɕyn ^h -tʃau ^h	att söka efter, sökes
11 地	11 里	43	dìlǐ	ti ^h -li ^h	på fältet
12 劳	12 动	24	láodòng	lau ^h -tuŋ ^h	rrbeta
13 新	13 奇	12	xīnqí	ɕin ^h -tɕ ^h i ^h	nyfiken
14 留	14 学	22	liúxué	liou ^h -ɕye ^h	att studera utomlands
15 实	15 际	24	shíjì	ʃi ^h -tɕi ^h	verklighet
16 体	16 验	34	tǐyàn	t ^h i ^h -jan ^h	uppleva
17 终	17 于	12	zhōngyú	tʃuŋ ^h -y ^h	äntligen
18 代	18 表	43	dàibiǎo	tai ^h -piau ^h	representera
19 凉	19 快	20	liángkuài	lian ^h -k ^h wai ^h	kyld
20 香	20 港	13	xiānggǎng	ɕiaŋ ^h -kaŋ ^h	Hong Kong
21 钱	21 包	21	qiánbāo	tɕ ^h ien ^h -pau ^h	plånbok
22 结	22 果	23	jiéguǒ	teie ^h -k ^h wuo ^h	resultat
23 托	23 运	14	tuōyùn	t ^h uo ^h -yn ^h	Frakt
24 箱	24 子	10	xiāngzi	ɕiaŋ ^h -tsi ^h	paket, bagage
25 卫	25 生	41	wèishēng	wei ^h -ʃəŋ ^h	Hygien

*I tabellen transkriberas orden med Pinyin och IPA (Lee & Zee 2003) och slutligen anges svensk betydelse.

Till vänster i tabellen återfinns de 25 orden i den ordning de hade i diktamen. Varje ord består av två stavelser, σ1 plus σ2, t.ex. ord nr. 1 黄河 <huánghé> där stavelsen 1 黄 <huáng> är den första och 1 河 <hé> den andra. Alla ord som presenterades är lexikala (till skillnad från de logatomer, betydelösa ord, som använts i andra undersökningar som t.ex. Chuang et al. 1972).

I kolumnen TK följer tonerna till dessa ord (tonkombinationer, TK, t.ex. innehåller <huánghé> två T2, T2+T2, som förkortas 22). Som framgår av tabell 1 finns följande TK inte representerade i materialet: T3+T2, T4+T2, T3+T0 och T4+T0.

Stavelsen <liáng> förekommer dels i ord nr. 2 凉爽 <liángshuǎng> och dels i ord nr. 19 凉快 <liángkuai>; stavelsen <xiāng> finns dels i ord nr. 20 香港 <xiānggǎng> och dels i ord nr. 24 箱子 <xiāngzi> och slutligen finner man stavelsen <kè> som utgör andra stavelsen dels i ord nr. 4 旅客 <lǚkè> och dels i ord nr. 6 立刻 <likè>; märk dock att i detta sista exempel föreligger inte homografi men väl både homofoni och homotonalitet.

I materialet förekommer några exempel på stavelser som är homofona men heterotonala, <zhaò> som finns dels i ord nr. 9 照相 <zhàoxiàng> och nr. 10 寻找 <xúnzhǎo>, stavelsen i ord nr. 6 立刻 <likè> och i ord nr. 11 地里 <dìlǐ>, samt <xiang> i ord nr. 9 照相 <zhàoxiàng> och ord nr. 20 香港 <xiānggǎng> (nr. 24 箱子 <xiāngzi>). T0 återfinns i två ord: stavelsen <zi> i ord nr. 24 箱子 <xiāngzi> respektive <kuai> i ord nr. 19 凉快 <liángkuai>. Konsonantkategorierna och tonkombinationerna i materialet ovanför är inte symmetriskt. Trots att det finns några stavelser som har gemensam kärna och koda kan <xiāng> och <liáng> enbart tas upp för att göra en jämförelse, men data är väldigt få.

Bilaga 2

Fördelningen av konsonantkategorierna av de 25 tvåstaviga orden

	Fonologiska särdrag					$\sigma 1$				$\sigma 2$				
	[son]	[app]	[lat]	[klu]	[asp]	T1	T2	T3	T4	T1	T2	T3	T4	T0
1. [l]	+	+	+				2, 12, 14, 19	4	6			11		
2. [w, ɹ, j]	+	+	–						25		17		8	16, 23
• [m, n, ŋ]	+	–	–											
•														
3. [p ^h]	–			+	+									
• [t ^h]	–			+	+	23								
• [ts ^h]	–			+	+									
• [tʃ ^h]	–			+	+									
• [te ^h]	–			+	+	3	21				13			
• [k ^h]	–			+	+	7		16					4, 6, 7	19
•														
4. [p]	–			+	–					5, 21		18		
• [t]	–			+	–				11, 18				12	
• [ts]	–			+	–									24
• [tʃ]	–			+	–	17			9			10		
• [te]	–			+	–		22						15	
• [k]	–			+	–		5					20, 22		
•														
5. [f]	–			–						3				
• [s]	–			–										
• [ʃ]	–			–			15			25		2		
• [e]	–			–		8, 13, 20, 24	10				14			
• [x]	–			–			1				1		9	
•						8	10	2	5	4	4	7	8	2
•														

* Bilaga 2 ovan hänvisar till ordningsföljden för testorden (jfr. bilaga 1). Spalterna $\sigma 1$ och $\sigma 2$ i tabellen visar stavelsens ton och segment. Ta som exempel nr. 2 i kolumnen T2 under $\sigma 1$ i tabellen: här kan tre fakta utläsas, dels att det gäller den första stavelsen <liáng> i testord nr. 2 凉爽 <liángshuǎng>, dels att stavelsen uttalas med T2 och dels att initialen utgörs av [l] som klassificeras [+son, +app, +lat]. I studien räknas affrikatorna som en subkategori under klusilerna (Maddieson & Disner 1984).

Bilaga 3

Fördelningen av totala antalet felmarkeringar inom σ1

σ1		Konsonantkategorier (KONK)																								Antal	
		[+son]						[-son]																			
		[+appr]						[+klu]												[-klu]							
		[+lat]						[-lat]						[+asp]						[-asp]							
S	Nr.	2	4	6	12	14	19	25	21	3	7	16	23	5	9	11	17	18	22	1	8	10	13	15	20	24	
	Pinyin	liáng	lǚ	lì	láo	liú	liáng	wèi	qián	qiū	kāi	tǐ	tuō	guǎng	zhào	dì	zhōng	dài	jié	huáng	xī	xún	xīn	shí	xiāng	xiāng	
	Tecken	凉	旅	立	劳	留	凉	卫	钱	秋	开	体	托	广	照	地	终	代	结	黄	吸	寻	新	实	香	箱	
Tonförväxlingsmönster (TFM)	T1→T'2									3	4		14								7		7		15	19	69
	T1→T'3									1			2								3				8	5	19
	T1→T'4									15	26		16				1				5		6		4	6	77 165
	T2→T'1	6			2	4	2		17									7	10		24		4			76	
	T2→T'3	8			39	1	22		1										2				4			77	
	T2→T'4	8			5	17	2		2									9	7		15		18			83 236	
	T3→T'1		2									3		10												15	
	T3→T'2		13									33		19												65	
	T3→T'4		8									2		12												22 102	
	T4→T'1			3				1								3	3									10	
T4→T'2			8				13							2	9		3								35		
T4→T'3			5											2											7 52		
Antal		22	23	16	46	22	26	14	19	18	30	37	32	41	7	12	1	3	16	19	15	40	13	26	27	30	555
		155						14	136						80						170						555

Bilaga 4

Fördelningen av totala antalet felmarkeringar inom σ2

σ2		Konsonantkategorier (KONK)																										Antal		
		[+son]													[-son]															
		[+appr]					[+klu]																		[-klu]					
		[+lat]	[-lat]				[+asp]					[-asp]																		
S	Nr.	11	8	16	17	23	4	6	7	13	19	5	10	12	15	18	20	21	22	24	1	2	3	9	25	14				
	Pinyin	li	yǐn	yàn	yú	yùn	kè	kè	kuò	qí	kuai	bō	zhǎo	dòng	jì	biǎo	gǎng	bāo	guǒ	zi	hé	shuǎng	fēng	xiàng	shēng	xué				
Tecken	里	引	验	于	运	客	刻	阔	奇	快		播	找	动	际	表	港	包	果	子	河	爽	风	相	生	学				
Tonförväxlingsmönster (TFM)	T1→T'2											15						3					2		3		23			
	T1→T'3											2						1									3			
	T1→T'4											3						9					3		9		24			
	T1→T'0											2						2							1	5	55			
	T2→T'1									2											1						3			
	T2→T'3				52					39											18					6	115			
	T2→T'4				2					3											1					3	9			
	T2→T'0									1																1	2	129		
	T3→T'1		2														1	7					1				11			
	T3→T'2		28										3				7	17		12			18				85			
	T3→T'4		3										1				2	1		1			2				10			
	T3→T'0																										106			
	T4→T'1			4		11			1	1				14	9										7		45			
	T4→T'2			13		5		3	3	7					1	3									5		36			
	T4→T'3			6				2	4	3					1										1		19			
T4→T'0			2		2				1					6	1									3		12	119			
T0→T'1																			3							3				
T0→T'2											1															1				
T0→T'3																										0				
T0→T'4										39										28						67	71			
Antal		0	33	25	54	18	5	8	12	45	40	22	4			10	25	15	13	31	20	21	5	16	13	10	480			
		0				130					110									155						85				

Bilaga 5

Förhållandet mellan konsonantkategorier (KONK) och tonförväxlingsmönster (TFM) inom $\sigma 1$

S	TFM	1. [+lat]				2. [-lat]				3. [+asp]				4. [-asp]				5. [-klu]				Σ			
		N	%	M	SD	N	%	M	SD	N	%	M	SD	N	%	M	SD	N	%	M	SD	N	%	M	SD
T1	→T'2									21	3,78							48	8,65			69	12,43		
	→T'3									3	0,54							16	2,88			19	3,42		
	→T'4									55	9,91			1	0,18			21	3,78			77	13,87		
												4,74	4,76			,60	,10			5,10	3,10	166		1,98	3,29
T2	→T'1	14	2,52							17	3,06			7	1,26			38	6,85			76	13,69		
	→T'3	70	12,61							1	0,18							6	1,08			77	13,87		
	→T'4	32	5,77							2	0,36			9	1,62			40	7,21			83	14,95		
				6,97	5,15							1,2	1,61			,96	,85			5,05	3,44	236		2,83	3,70
T3	→T'1	2	0,36							3	0,54			10	1,80							15	2,70		
	→T'2	13	2,34							32	5,77			19	3,42			1	0,18			65	11,71		
	→T'4	8	1,44							2	0,36			12	2,16							22	3,96		
				1,38	,99							2,22	3,07			2,46	,85			,06	,10			1,22	1,66
T4	→T'1	3	0,54			1	0,18							6	1,08							10	1,80		
	→T'2	8	1,44			13	2,34							14	2,52							35	6,31		
	→T'1	5	0,90											2	0,36							7	1,26		
				,96	,45			,84	1,30							1,32	1,10							,62	,87
		155	27,93			14	2,52			136	24,50			80	14,41			170	30,63			555	100		

Bilaga 6

Förhållandet mellan konsonantkategorier (KONK) och tonförväxlingsmönster (TFM) inom $\sigma 2$

S	TFM	1. [+lat]				2. [-lat]				3. [+asp]				4. [-asp]				5. [-klu]				Σ			
		N	%	M	SD	N	%	M	SD	N	%	M	SD	N	%	M	SD	N	%	M	SD	N	%	M	SD
T1	→T'2													18	3,75			5	1,04			23	4,79		
	→T'3													3	0,63							3	0,63		
	→T'4													12	2,50			12	2,50			24	5,00		
	→T'0													4	0,83			1	0,21			5	1,04		
T2	→T'1									2	0,42														
	→T'3					52	10,83			39	8,13							24	5,00			115	23,96		
	→T'4					2	0,42			3	0,63							4	0,83			9	1,88		
	→T'0									1	0,21							1	0,21			2	0,42		
T3	→T'1							2,81	5,35			2,35	3,86							1,56	2,31			1,34	3,02
	→T'2					2	0,42							8	1,67			1	0,21			11	2,29		
	→T'4	1	0,21			28	5,83							39	8,13			18	3,75			85	17,71		
	→T'0					3	0,63							5	1,04			2	0,42			11	2,29		
T4	→T'1							1,72	2,75							2,71	3,68			1,10	1,78			1,12	2,22
	→T'2					15	3,13			2	0,42			23	4,79			7	1,46			47	9,79		
	→T'3					18	3,75			13	2,71			4	0,83			5	1,04			40	8,33		
	→T'0					6	1,25			9	1,88							1	0,21			16	3,33		
T0	→T'1					4	0,83			1	0,21			5	1,04			3	0,63			13	2,71		
	→T'2							2,24	1,42			1,30	1,10			1,67	2,13			,84	,54			1,21	1,38
	→T'4	2	0,42							39	8,13			28	5,83							69	14,38		
	→T'0											2,09	4,03			1,62	2,83							,76	2,17
		3	0,63			130	27,08			110	22,92			152	31,67			85	17,71			480	100		

Bilaga 7

Tonkombinationstimuli (TKS) × Felaktiga TonKombinationsResponder (FTKR)

TKS	FTKR																			
	1+1		1+2		1+3		1+4		1+0		2+1		2+2		2+3		2+4		2+0	
	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%
1+1							3	,29			2	,19					2	,19	1	,10
1+2	2	,19			80	7,73	4	,39	1	,10			3	,29	8	,77			1	,10
1+3	1	,10	29	2,80			1	,10			4	,39	8	,77	14	1,35	4	,39	12	1,16
1+4	1	,10	5	,48					3	,29	6	,58	2	,19	2	,19	12	1,16	2	,19
1+0							22	2,13			4	,39					20	1,93	7	,68
2+1	12	1,16			2	,19	8	,77					2	,19			6	,58	2	,19
2+2			11	1,06	6	,58									14	1,35	3	,29	1	,10
2+3			14	1,35	29	2,80	2	,19					16	1,55			3	,29	2	,19
2+4	6	,58					3	,29			4	,39	1	,10	1	,10	1	,10	14	1,35
2+0									2	,19							22	2,13		
3+1	4	,39	10	,97					2	,19	15	1,45	4	,39			2	,19		
3+4							4	,39			8	,77	14	1,35	11	1,06	29	2,80	2	,19
4+1									2	,19	9	,87	2	,19			6	,58		
4+3									3	,29					2	,19	2	,19	9	,87
4+4	2	,19			2	,19	3	,29	2	,19	2	,19					9	,87		