

韓国語ソウル方言における語中閉鎖音の知覚

邊 姫京 (国際教養大学)

byun@aiu.ac.jp

1. はじめに

邊 (2017) はソウル方言における語中閉鎖音の音響特徴を明らかにするために関連が指摘されているパラメータのうち、VOT、閉鎖区間長 (CD)、全長 (TD=VOT+CD)、先行母音長、後続母音の fo について検討した。それによれば、いずれのパラメータも単独では3種の閉鎖音を区別することができないが、VOT、CD、TDは3種の閉鎖音を2つのグループに区別することができる。そしてVOTとCD、あるいはVOTとTDのように2つのパラメータを組み合わせることで3種の閉鎖音をそれぞれのカテゴリーに分けることができる。具体的には、図1で見ると、激音とそれ以外の子音 (濃音・平音) はVOTの違いで、濃音と平音はCDまたはTDの違いで分けられる。つまり、激音は長いVOT、濃音は短いVOTと長いCD (またはTD)、平音は短いVOTと短いCD (またはTD) に特徴づけられる。

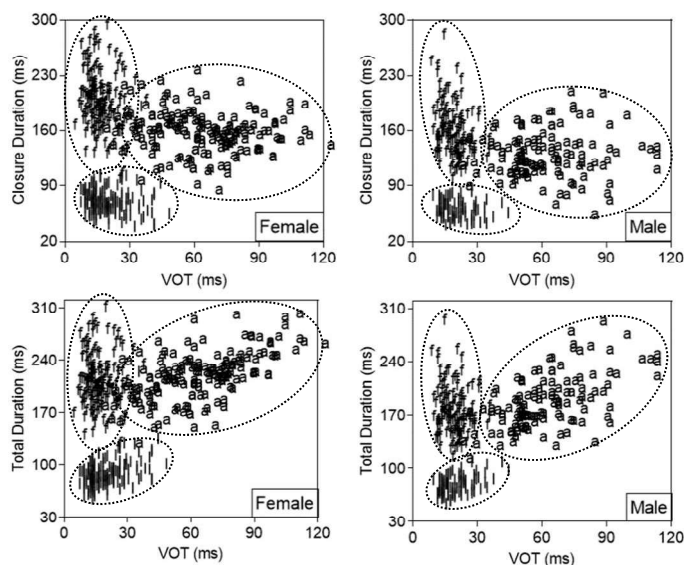


図 1: 語中閉鎖音の生成におけるカテゴリー域 (邊 (2017) の図 8)

横軸: VOT, 縦軸: CD (上段) と TD (下段), 図中の a: 激音 /tʰ/ /kʰ/, l: 平音: /t/ /k/, f: 濃音 /t*/ /k*/

本研究は、生成における音響特徴が知覚においても有効であるかを調べたものである。VOT と CD に焦点を当て、2つのパラメータを段階的に変化させた合成音を作成し、ソウル及び京畿道出身の韓国語母語話者を対象に聴取実験を行った。以下にその詳細を報告する。

2. 研究方法

2.1. 刺激音

原音は、1987年生まれのソウル方言話者（男性）が発話した濃音の「ak*a (아까)」である。原音をどの子音にするかは先行研究でもまちまちであるが、Oh (2019) は、激音と平音の場合、どちらを使っても結果には影響しないことを報告している。本研究では元の音色を変えずに操作できたのが濃音だったので濃音を使用した。VOT の操作には Praat の manipulate 機能を利用したが、使える音声のうち激音は圧縮によりバーストの部分が強くなり過ぎて平音の音色が出にくく、平音は濃音と激音に感じられる緊張が感じられないので使用しなかった。軟口蓋音にしたのは、VOT 操作のためにある程度の長さの VOT が必要だったからである。

VOT と CD の幅は、邊 (2017) の結果を基に、それぞれ 5~85ms, 40~240ms に決めた。図 2 は 1990 年代生まれの男性 (M) と女性 (F) の発話結果である。図 1 は 1953~1999 年生まれの話者の結果であるが、今回の聴取実験に参加する参加者を 1990 年代以降生まれの若者に限定したので刺激音の幅も 1990 年代生まれの話者の発話を参照した。原音の VOT は 25ms, CD は 137ms である。CD は 140ms にしてから操作を開始した。まずは VOT を 10ms 間隔で 9 段階に変化させ、その後 CD を 20ms 間隔で 11 段階に変化させた。

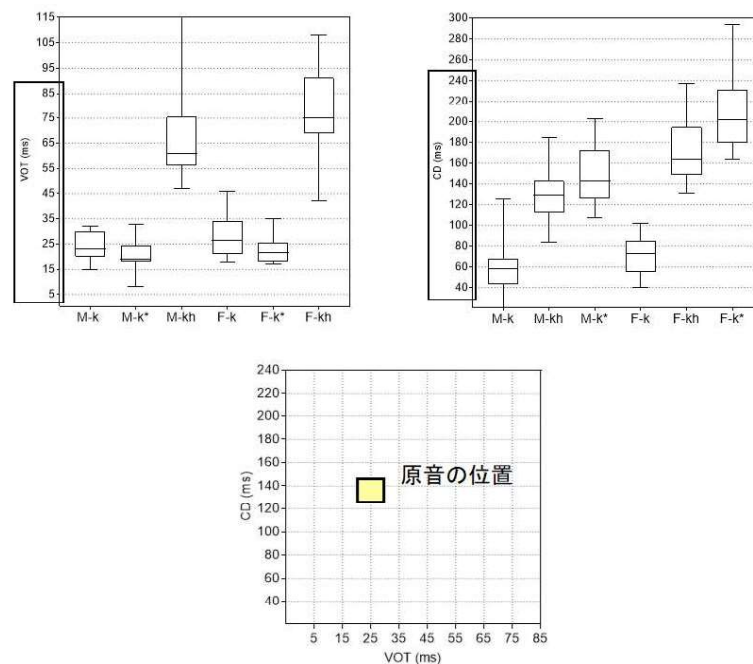


図 2: 軟口蓋音/ak*a/の測定値(1990 年代生まれのソウル方言話者, 邊(2017)から作成)

2.2. 参加者 (聴者)

ソウル及び京畿道で生まれ育った 1990~2001 年生まれの男女 34 名で、全員大学生である。

2.3. 手順

実験は大学の研究室で、Praat 上で行った。参加者はヘッドホンから流れる音声を聴いてもっとも近いと思われるものをパソコン画面上に表示されるハングル表記の「aka」「ak^ha」「ak*^a」からクリックして選ぶ。刺激音は必要であればさらに2回まで聴き直すことができる。回答を間違えた場合は再回答ができるように設定した。本実験に先立って6問を使って本実験と同じ形式で練習を行った。本実験は100問（刺激音99問とダミー1問）で、25問ごとに短い休憩がある。回答に制限時間は設けず参加者の裁量に任せたが、かかった時間は長い人でも5分以内であった。事前のアンケートで聴力に問題があると報告した参加者はいなかった。参加者には実験後に謝礼を渡した。

3. 結果

3.1. パラメータ別結果

図3にVOT、図4にCDの結果を示す。図3で見るとVOTが長くなるほど刺激音の同定率が上がり、逆にVOTが短いほど濃音あるいは平音の同定率が上がる。濃音と平音は35ms以上では両者の間に差がほとんどない。両者の同定率に差が出るのは25ms以下で、濃音の同定率は平音のそれより倍以上高い。刺激音と濃音は30msを境にきれいに分かれており、VOTが刺激音と濃音を区別する主要なパラメータであることがわかる。つまり、長いVOTであれば刺激音、短いVOTであれば濃音に優先的に同定されると言える。ただし、刺激音がVOTのみで9割近くまで同定されるのに対して、濃音はもっとも短いVOTでも同定率は7割程度で、VOTだけで濃音が同定できるとは限らない（図3では15ms以下で同定率にほとんど変化がない）。濃音と同様に短いVOTに反応する平音ももっとも短いVOTでの同定率は3割程度で、VOTだけの同定は濃音よりもさらに精度が落ちる。

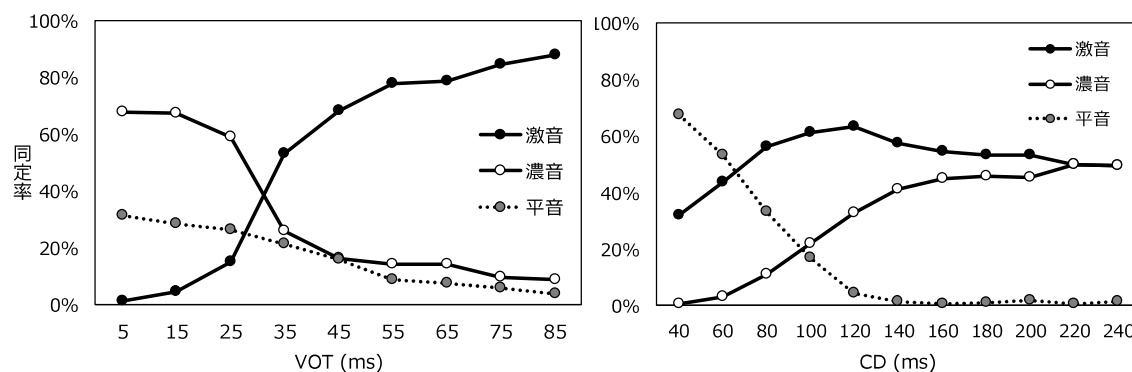


図 3: VOT の同定率

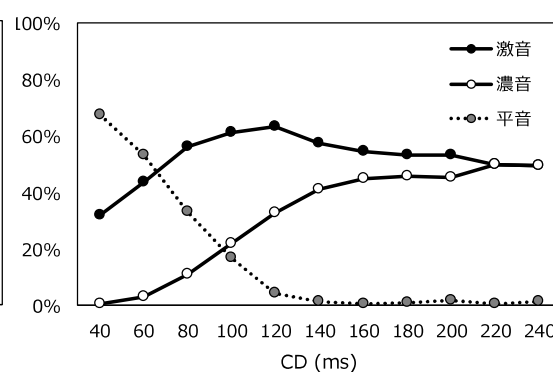


図 4: CD(閉鎖区間長)の同定率

図4のCDの結果を見よう。平音は100ms以下であれば、CDが短いほど同定率が上がる。図4でもっとも短い値は40msであるが、40ms時点でも平音のグラフは上向きになっており、さらに短い値であれば同定率はさらに上がることが予想される。このCDの結果と前述のVOTの結果を合わせると、平音の同定にかかわる主要なパラメータはCDであり、

VOT も手がかりになると言えそうである。濃音は CD が長いほど同定率が上がる傾向にあるが、これが言えるのはおよそ 160ms までで、それ以上は長くなっても大きな変化はない。激音は 80ms 以上では同定率にあまり差がなく、CD の効果は VOT に比べて低いと言えよう。

図 3, 4 の結果を整理すると、激音の知覚には主に長い VOT が用いられる。濃音の知覚には短い VOT と長い CD が用いられる。平音の知覚には短い VOT と短い CD が用いられる。

3.2. VOT と CD の組み合わせの結果

図 5, 6 に VOT と CD を組み合わせた 3 次元グラフを示す。横軸は VOT (ms), 縦軸は CD (ms), 各セルの濃さは同定率を表す。図 5 は各子音の同定率を 20% 間隔で示した。図 6 は当該子音の同定率で、子音の記載のない 50% のセルは、隣接する子音と同定率が 50% で同値の場合である。

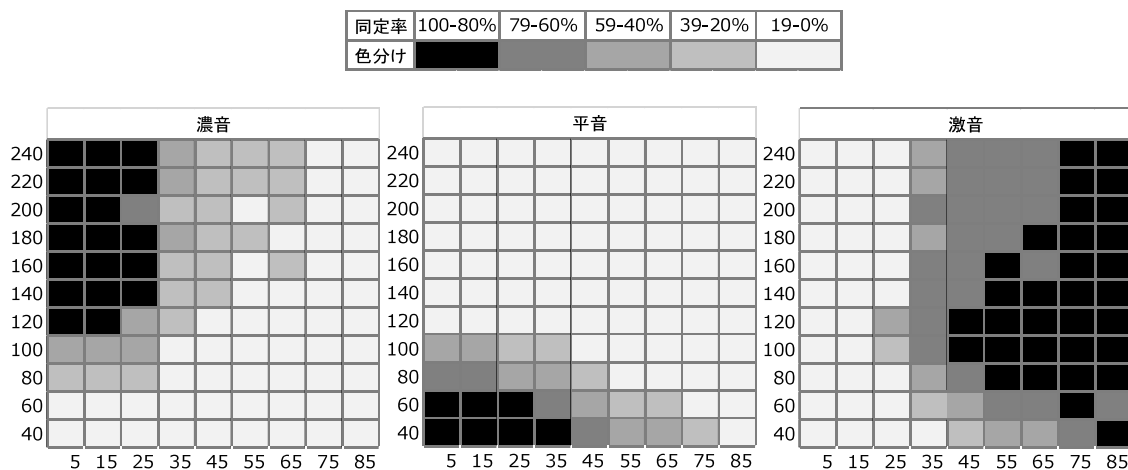


図 5: 各子音の同定率

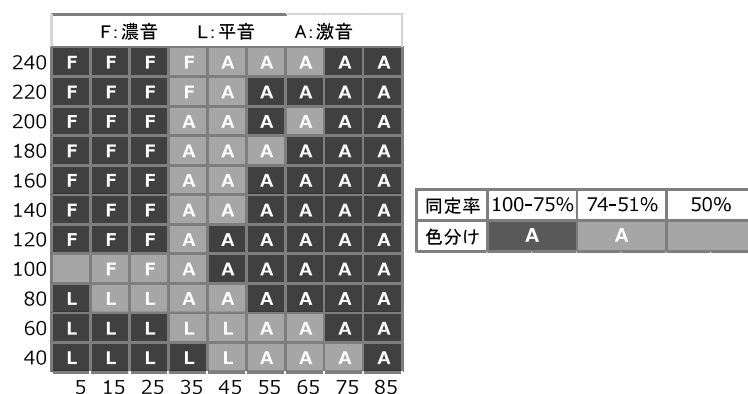


図 6: 3 種の閉鎖音の同定率(語中閉鎖音の知覚におけるカテゴリー一域)

図 5 で見るように、濃音は短い VOT と長い CD、平音は短い VOT と短い CD、激音は概して 55ms 以上であれば CD と関係なく VOT のみで知覚されている。図 6 では各子音の

テゴリー域の相対的な位置関係を見ることが出来る。VOT の違いで激音とそれ以外の子音（濃音と平音）、CD の違いで濃音と平音が区別されており、これは図 1 で見た生成におけるカテゴリー域と一致する。これにより語中の 3 種の閉鎖音は、生成、知覚のいずれも VOT と CD の違いで区別されると言える。

3.3. パラメータの影響度

同定における各パラメータの相対的な影響力を検討するために多項ロジスティック回帰分析を行った (RStudio 使用 Version 1.1.456)。分析には R (R Core Team 2016) のパッケージ `nnet` の `multinom` 関数を用いた。説明変数は VOT と CD、目的変数は子音種（激音、濃音、平音の 3 択）である。VOT と CD は中心化を行った。参照カテゴリーは激音である。

表 1 に回帰分析の結果、表 2 にオッズ比と信頼区間を示す。表 1 の係数 β (推定値) は、プラスの値であれば参照カテゴリーである激音に対して濃音または平音に同定される確率が高く、マイナスの値であれば参照カテゴリーである激音に同定される確率が高いことを表す。また、プラス、マイナスの符合を除いた値 (絶対値) は目的変数への影響の大きさを表す。具体的には、濃音の CD の係数はプラスなので濃音に同定される確率が高く、濃音の VOT と平音の VOT、CD の係数はマイナスなので参照カテゴリーである激音に同定される確率が高い。係数の絶対値からは、激音と濃音、激音と平音の知覚には CD よりも VOT の影響力が強く (激音 vs 濃音: VOT の $|-0.073| > CD$ の $|0.011|$, 激音 vs 平音: VOT の $|-0.077| > CD$ の $|-0.047|$)、濃音と平音の知覚には VOT よりも CD の影響力が強いことがわかる (濃音 vs 平音: CD の $|0.011 - (-0.047)| > VOT$ の $|(-0.073) - (-0.077)|$)。なお、VOT と CD の P 値はいずれも有意である。

表 1: 多項ロジスティック回帰分析の結果

	濃音 (激音 vs 濃音)				平音 (激音 vs 平音)			
	係数 β	標準誤差	z 値	P 値	係数 β	標準誤差	z 値	P 値
切片	-0.843	0.060	-14.141	<.0001	-3.381	0.174	-19.405	<.0001
VOT	-0.073	0.003	-27.611	<.0001	-0.077	0.004	-21.283	<.0001
CD	0.011	0.001	11.077	<.0001	-0.047	0.002	-19.175	<.0001

表 2: オッズ比と信頼区間

	濃音 (激音 vs 濃音)			平音 (激音 vs 平音)		
	オッズ比	2.5%	97.5%	オッズ比	2.5%	97.5%
切片	0.430	0.383	0.484	0.034	0.024	0.048
VOT	0.930	0.925	0.934	0.926	0.919	0.932
CD	1.011	1.009	1.013	0.955	0.950	0.959

では、具体的にどのくらい影響力が強いかは、表2のオッズ比(推定値)より推定することができる。オッズ比は1を基準に1より大きい場合は当該のカテゴリー、1より小さい場合は参照カテゴリーに同定される確率を表す。つまり、濃音のCDのオッズ比は1より大きい1.011なので、CDが長いほど激音に対して濃音に同定される確率が1.011倍高い。また、オッズ比が1より小さい濃音のVOTと平音のVOT、CDは、VOTまたはCDが長いほど参照カテゴリーである激音に同定される確率がそれぞれ0.93倍、0.926倍、0.955倍高いことになる。オッズ比は、説明変数が連続変数の場合は1単位増すごとの増加分になるので、激音と濃音ではCDが1ms増すごとに激音より濃音に同定される確率が約1%増すと解釈できる。VOTは1ms増すごとに濃音より激音に同定される確率が約8%増す(激音に対して濃音に同定される確率である0.930の逆数、 $1/0.93 \div 1.075$)。同様に、VOTとCDが1ms増すごとに平音に対して激音に同定される確率はそれぞれ約8%と5%増す(VOT: $1/0.926 \div 1.079$, CD: $1/0.955 \div 1.047$)。濃音と平音の場合、VOTとCDが1ms増すごとに平音に対して濃音に同定される確率はそれぞれ約0.4%と6%になる(VOT: $0.930/0.926 \div 1.004$, CD: $1.011/0.955 \div 1.058$)。

4. まとめ

以上、生成における音響特徴が知覚においても有効であるかを、VOTとCDを対象に検討した。結果で明らかになったように、VOTとCDは生成と同様に知覚においても有効なパラメータであり、カテゴリー間の相対的な位置関係も生成の場合とほぼ同様である。以下に結果をまとめる。

語中における3種の閉鎖音は、VOTとCDの組み合わせで、長いVOTであれば激音、短いVOTと長いCDであれば濃音、短いVOTと短いCDであれば平音と知覚される。激音と濃音、激音と平音におけるVOTの境界はともに30ms前後、激音と平音、濃音と平音におけるCDの境界はそれぞれ60ms前後、100ms前後であった。この境界は軟口蓋音の場合で、軟口蓋音よりVOTが短い両唇音、歯茎音では異なり得る。多項ロジスティック回帰分析では係数 β の比較とオッズ比の比較から各パラメータの相対的な影響力について議論した。

本研究は、JSPS KAKENHI 17K02685の助成を受けたものである。

参考文献

- 邊姫京 (2017) 「韓国語ソウル方言における語中閉鎖音の音響特徴」『音声研究』 21:2, 61-79.
- Oh, Eunjin. (2019) "The effect of base token selection for stimuli manipulation on the perception of lenis and aspirated stops in Seoul Korean." *Proceedings of the 2019 Spring Conference of Korean Society of Speech Sciences*, 79.
- R Core Team (2016) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.