

SPEECH-RATE AND TEMPO EFFECTS ON THE ORGANIZATION OF SPOKEN AND SUNG SYLLABLES

Nicole SCOTTO DI CARLO

LPL - UMR 6057 CNRS - France
nicole.scotto@lpl.univ-aix.fr

ABSTRACT

In singing, syllables undergo considerable lengthening. This modifies their temporal structure by increasing the proportion of vowels and decreasing the proportion of consonants. Vowel lengthening is due both to the substantial slowing of the tempo inherent in all sung productions, and to longer articulation times brought about by the large extent of the movements made by the phonatory organs. Consonant shortening, on the other hand, is the result of physiological and aesthetic constraints.

Except for liquid and nasal consonants, whose acoustic structure is similar to that of vowels, the elasticity of sung consonants is practically identical to that of spoken ones because of the physiological constraints to which they are subjected.

The expansion of vowels is limited at the physiological level only by the singer's breath length, except in the case of nasal vowels which cannot be held for longer than a few hundredths of a second due to the inertia of the velum. The vowel compression limit is very close to that recorded for the spoken voice, most likely for perceptual reasons.

Keywords: Syllabic organization, speech rate, tempo, phonemic plasticity, singing voice.

1. INTRODUCTION

After a review of the main hypotheses and theories set forth to explain the greater duration of vowels in vocal emissions produced under extreme conditions, the effects of speech rate and tempo on the relative durations of vowels and consonants in spoken and sung syllables will be compared. Then, following an evaluation of phonemic elasticity, the elasticity limits of phonemes in speech and singing will be considered.

2. EXPERIMENTAL SETUP

In order to work in an experimental situation as similar as possible to the conditions in which lyrical artists are accustomed to singing, the text used here was not a nonsense sequence but an excerpt from an opera sung in French and selected so that every phoneme in the language was represented. For this pilot study, a professional soprano read the text normally (NSp = normal speaking voice), as slowly as possible (SSp = slow speaking voice), and as rapidly as possible (FSp = fast speaking voice). Then she was asked to sing the corresponding excerpt at the tempo recommended by the composer and set on a metronome ♩ = 80 (NSi = normal singing voice), at half that tempo ♩ = 40 (SSi = slow singing voice), and at twice that tempo ♩ = 160 (FSi = fast singing voice). She was also asked to pronounce a 24-syllable sequence in her usual speaking manner, and then to sing the sequence on A#₃ (233 Hz) for the lower register, A#₄ (466 Hz) for the middle register, and A#₅ (932 Hz) for the upper register, at the same rate as in her speaking voice. All recordings were made in an anechoic chamber, transcribed phonetically, and input into an ATFoIC acoustic analysis (amplitude, time, fundamental frequency, and intensity curve).

3. PHONEME DURATION IN EXTREME VOCAL PRODUCTIONS

3.1 Hypotheses

The French language is characterized by its syllabic isochrony, a phonetic law whereby spoken syllables tend to be of equal duration. Isochrony cannot be achieved in singing, where the mean duration of vowels is much longer and the mean duration of consonants is slightly shorter than in speech (Scotto di Carlo, 2005). This phenomenon is also found to a certain extent in the shouting voice, where vowel

duration rises by 50% and consonant duration drops by 20% (Rostolland, 1979). Some authors ascribe the phenomenon to vocal effort (Rostolland, 1982; Bonnot & Chevrie-Muller, 1991; Fónagy & Fónagy, 1966).

It is difficult to rely on the notion of vocal effort, however. Not only does this concept have different meanings for different authors, but the traditional way of measuring it by the Sound Pressure Level (SPL), has recently been questioned. Several Swedish studies showed that the SPL does not play a major role in the perception of vocal effort (Rundlöf, 1996; Traunmüller, 1997; Eriksson & Traunmüller, 1999). This finding led to the development of a more reliable method for measuring vocal effort which takes its perceptual aspect into account by incorporating a proxemic factor (Traunmüller & Eriksson, 2000). Regarding the singing voice, Kallopoulos (2000) hypothesized that the longer duration of vowels and the shorter duration of consonants is linked to articulatory force. Yet according to a study conducted by Scotto di Carlo (1979), the reverse phenomenon takes place when the articulatory force increases, since an overarticulated singing voice exhibits a slight decline in vowel duration (14%) accompanied by a sharp rise in consonant duration (58%). Moreover, due to opening and aperture constraints, the articulatory force is not as great in singing as in speech, as shown in palatographic studies on professional singers, where tongue contact with the hard palate gets increasingly weak and small as the singer goes from the lower to the middle register, ending up barely touching it in the upper register (Scotto di Carlo, 1994).

In addition to these hypotheses — which explain the syllabic imbalance characteristic of extreme vocal productions in terms of vocal effort or articulatory force — it is useful to consider an observation that concerns only the singing voice and introduces the notion of aesthetic constraint. In singing, consonant shortening is all the more marked due to the fact that, unlike vowels which provide an ideal melodic medium, consonants interrupt the sound flow. This leads singers to reduce consonantal durations in an attempt to preserve the quality of the *legato*¹ (Scotto di Carlo, 1979). Consonant shortening also depends on the singer's articulatory strategy. It is greater for higher voices, which fall in a range of frequencies where articulation is difficult (Scotto di Carlo, 1994).

3.2 Theories

In a comparative study of the spoken and shouted voices, Schulman (1989) recalls the various interpretations given to vowel lengthening:

- In the *mechanistic theory*, the velocity of the articulating organs cannot increase in proportion to the extent of the required movements, so more time is needed to complete them (Lindblom, 1967).

- In the *communicative theory*, vowels are the sounds of the language that are the most audible from a distance, even in very noisy environments. As such, they convey the most information, so speakers intentionally lengthen them when shouting (Fónagy & Fónagy, 1966; Clark *et al.*, 1987).

Lindblom's mechanistic theory is the most suitable one for explaining the phenomena observed in the singing voice, where the movements of the phonatory organs cover a much greater distance than in the spoken and shouted voices, and for this reason, take more time to achieve. This was demonstrated here in the sequences spoken and sung at the same syllable rate, where the total singing duration exceeded the total speaking duration by 6.3% in the lower register, 13 % in the middle register, and 28.7% in the upper register.

Although considerable vowel lengthening in extreme production conditions has been observed by all authors, the duration of consonants under such conditions remains a controversial topic. This is most likely due to the fact that phenomena linked to vocal effort are difficult to distinguish from phenomena related to articulatory force. These two processes may be dissociated as they are in singing, where they do not have the same temporal effects on phonemes (Scotto di Carlo, 1979).

4. EFFECTS OF SPEECH RATE AND TEMPO ON SYLLABLE BALANCE

The mean durations of both consonants and vowels increased when the speech rate or tempo slowed down. However, while the vowel durations rose considerably, the range of variation in the consonantal durations was limited.

Modifying the speech rate or tempo caused changes in the syllable duration, which had the effect of altering the consonant/vowel balance within the syllable. The slower the speech rate

or tempo, the longer the syllable duration and the smaller the proportion of consonants.

4.1 Vowel Elasticity

The controlled study of speech rate showed that the high elasticity of vowels caused them to exhibit greater temporal variations than consonants. One exception was the fast speaking voice (FSp), where the consonants were 41% shorter than in the normal speaking voice (NSp) while the vowels were only 37% shorter. A possible explanation for the greater shortening of consonants is that consonants produced at such a rapid speech rate are necessarily underarticulated. Indeed, the speech rate of this soprano was 6.65 syllables/second in FSp versus 3.89 syllables/second in NSp.

For the singing voice, the vowels always exhibited greater temporal variations than the consonants, as in the speaking voice. The expansion rates were 76% for vowels and 38% for consonants. Concerning the compression rate, it was only 41% for vowels and 23% for consonants. The mechanistic theory can account for the difficulties experienced by the soprano when singing rapidly, where her syllable rate was only 3.26 syllables/second. Indeed, this singer was unable to sufficiently increase the tempo so as to reach the target rate of 4.12 syllables/sec, which in this particular case, corresponds to an 80% compression of syllable duration.

4.2 Consonant Elasticity

While the vowels adapted easily to the speech rate and tempo changes via compression or expansion, the consonants could not be compressed or expanded to as great an extent due to their strong physiological constraints. The degree of elasticity varied across consonants. In speech and in singing alike, the voiced stops [b], [d], and [g], which are subject to the strongest aerodynamic constraints, had the lowest elasticity levels (39 to 135 ms for speech and 39 to 78 ms for singing). The 39 ms value shared by the spoken and singing voices seems to be the low-end cutoff point for voiced-stop compression, even in the fast singing voice where the shortest observed duration was 38.85 ms. The vocalic consonants [l], [r], and [R], the least constrained, had the highest elasticity levels (34 to 224 ms in the speaking voice and 18 to 322 ms in the singing voice), making them the privileged phonemes in singing.

The consonants with the highest elasticity are generally the easiest ones to sing, because they are not subjected to heavy physiological constraints. This is the case of the liquids [l] and [r] and the nasals [m] and [n], for example, whose acoustic structure close to vowels facilitates cantillation. This may be why we spontaneously use these phonemes to sing a song whose words we do not know.

5. ELASTICITY LIMITS OF CONSONANTS AND VOWELS

The elasticity limits of the consonants were practically the same in the speaking and singing voices, with a compression limit of 16 ms for both speech and singing, and an expansion limit of 324 ms for the speaking voice and 340 ms for the singing voice.

5.1 Expansion and Compression of Consonants

The expansion of certain consonants, such as the voiced stops, is limited by physiological constraints of an aerodynamic nature. Indeed, the complete closing of the vocal tract characteristic of the production of this type of consonant momentarily raises the supraglottal pressure until it reaches its maximum during the consonantal hold. Because variations in the subglottal pressure are small, the increase in the supraglottal pressure equalizes the pressure levels above and below the glottis. At this moment, the amount of air passing through the glottis drops abruptly and tends toward zero. This decreased transglottal airflow rate provokes a decrease in the suction effect (Bernoulli effect), which is the principal driver of the vocal-cord vibration regime. In short, the longer a voiced stop is held, the greater the supraglottal pressure, the lower the transglottal airflow rate, and the more the Bernoulli effect is inhibited. This eventually stops the laryngeal vibration and leads to devoicing.

The compression limits of consonants seem to be linked to physiological constraints rooted in the inertia of the articulatory organs and/or the mode of production. For example, voiceless fricatives are produced with a large glottal opening that prevents the vocal cords from vibrating, and with extensive narrowing of the vocal tract that generates the turbulent airflow responsible for the friction noises characteristic of these consonants. However, the turbulent regime takes a certain amount of time to set in,

firstly, because the supraglottal pressure, which is very low due to the glottal opening, cannot rise until the vocal tract is partially closed by the constriction, and secondly, because the constriction and the transglottal airflow rate must be finely adjusted to permit the airflow — which used to be laminar — to become turbulent downstream of the constriction point. It is not surprising, then, that voiceless fricatives cannot be compressed without losing their identity.

5.2. Expansion and Compression of Vowels

Vowel expansion is great in singing, where it is limited musically only by the musical score and physiologically only by the singer's breath length. The only exception is nasal vowels, which are impossible to hold during singing because the inertia of the velum causes it to collapse after a few hundredths of a second onto the back of the tongue and obstruct the buccal cavity. To prevent the nasal vowel from becoming a sound void of buccal airflow — whose hypernasalized timbre is particularly unpleasant to the ear — singers use a stratagem consisting of holding the corresponding oral vowel without nasalizing until the very end.

The observed compression limits of the vowels were also very close in the speaking and singing voices (38 ms and 40 ms, respectively). The reasons for these compressibility limits are perceptual in nature, insofar as overly compressed phonemes lose their intelligibility.

6. CONCLUSION

In singing, substantial syllable lengthening modifies the relative durations of consonants and vowels within a syllable. The resulting effect is an increase in the proportion of vowels and a decrease in the proportion of consonants. The mean duration of both consonants and vowels increases when the speech rate or tempo slows down, but while vocalic duration rises considerably, consonantal duration varies within a limited range. Finally, due to the heavy physiological constraints to which they are subjected, sung consonants retain an intrinsic duration and degree of elasticity very close to their spoken counterparts.

Concerning the two temporal criteria that characterize the singing voice, the first, vowel lengthening, is due to the extensive slowing of the tempo inherent in all sung productions, together with greater articulation times linked to the larger extent of the movements made by the phonatory organs. The second, consonant

shortening, results not only from the powerful aerodynamic constraints that govern consonant production, but also from the aesthetic constraints singers impose upon themselves in order to achieve a proper *legato*, one of the essential dimensions of musicality.

7. REFERENCES

- [1] Bonnot, J.F., Chevré-Muller, C. 1991. Some Effects of Shouted and Whispered Conditions on Temporal Organization. *J. Phonetics* 19: 473-483.
- [2] Clark, J., Lubker, J., Hunnicutt, S. 1987. Some Preliminary Evidence for Phonetic Adjustment Strategies in Communication Difficulty. In Steele, R., Threadgold, T. (eds), *Essay in Honour of Michael Halliday*, Amsterdam: Benjamins.
- [3] Eriksson, A., Traunmüller, H. 1999. Perception of Vocal Effort and Speaker Distance on the Basis of Vowel Utterances. *Proc. 14th ICPHS*. San Francisco, 2469-2472.
- [4] Fonágy, I., Fonágy, J. 1966. Sound Pressure Level and Duration. *Phonetica* 15: 14-21.
- [5] Kallopioulos, N. 2000 *Syllable Temporal Patterns*. Doctoral dissertation. National and Capodistrian University of Athens.
- [6] Lindblom, B. 1967. Vowel Duration and a Model of Lip Mandible Coordination. *STL-QPSR*, 4: 1-29.
- [7] Rostolland, D. 1979. *L'audition de la parole en présence de bruit*. Doctoral dissertation. Université de Paris.
- [8] Rostolland, D. 1982. Acoustic Features of Shouted Voice. *Acustica* 50: 118-125.
- [9] Rundlöf, J. 1996. *Perceptuella letrådar vid auditiv bedömning av avståndet mellan talare och lyssnare*. D-Uppsats, Inst. för lingvistik. Stockholm universitet.
- [10] Schulman, R. 1989. Articulatory Dynamics of Loud and Normal Speech. *J. Acoust. Soc. Am.* 85, 1: 295-312.
- [11] Scotto di Carlo, N. 1979. Perturbing effects of Overarticulation in Singing. *J. Res. Singing* II, 2: 10-27.
- [12] Scotto di Carlo, N. 1994. L'intelligibilité de la voix chantée. *Médecine des Arts* 10: 2-15.
- [13] Scotto di Carlo, N. 2005. Structure temporelle de la syllabe parlée et chantée. *Médecine des Arts* 52: 2-11.
- [14] Traunmüller, H. 1997. Perception of Speaker Sex, Age and Vocal Effort. *Phonum* 4: 183-186.
- [15] Traunmüller, H., Eriksson, A. 2000. Acoustic Effects of Variation in Vocal Effort by Men, Women and Children. *J. Acoust. Soc. Am.* 107, 6: 3438 - 3451.

-
1. Art of singing by tying notes together

EFFETS DU DEBIT ET DU TEMPO SUR L'ORGANISATION DE LA SYLLABE PARLEE ET CHANTEE

Nicole SCOTTO DI CARLO

LPL - UMR 6057 CNRS - France
nicole.scotto@lpl.univ-aix.fr

RESUME

Dans le chant, la syllabe subit un allongement très important qui entraîne une modification de sa structure temporelle et a pour effet d'augmenter la proportion de voyelles et de diminuer la proportion de consonnes. L'allongement des voyelles est dû à l'important ralentissement du tempo inhérent à toute production chantée ainsi qu'à l'augmentation de la durée articulatoire liée à l'amplitude considérable des mouvements effectués par les organes phonateurs. En ce qui concerne l'abrègement des consonnes, il résulte de contraintes physiologiques et esthétiques.

Exception faite des consonnes liquides et nasales dont la structure acoustique est proche de celles des voyelles, l'élasticité des consonnes chantées est pratiquement identique à celle des consonnes parlées à cause des contraintes physiologiques auxquelles elles sont soumises. L'expansion des voyelles n'est limitée sur le plan physiologique que par la tenue de souffle du chanteur, à l'exception des voyelles nasales que l'on ne peut tenir plus de quelques centièmes de secondes en raison de l'inertie du voile du palais. Quant à leur limite de compression, elle est très proche de celle enregistrée pour la voix parlée, vraisemblablement pour des raisons d'ordre perceptif.

Mots-clefs : Organisation syllabique, débit, tempo, plasticité phonémique, voix chantée.

1. INTRODUCTION

Après avoir passé en revue les principales hypothèses et théories sur l'accroissement de la durée des voyelles au cours de productions vocales réalisées dans des conditions extrêmes, nous comparerons les effets du débit et du tempo sur la durée relative des voyelles et des consonnes dans la syllabe parlée et chantée. Nous évaluerons ensuite l'élasticité des phonèmes avant d'en comparer les limites dans la parole et dans le chant.

2. PROTOCOLE EXPERIMENTAL

Afin de travailler dans une situation expérimentale qui se rapproche le plus possible des conditions dans lesquelles les artistes lyriques ont l'habitude de se produire, nous avons choisi non pas des logatomes mais un extrait d'opéra chanté en français et sélectionné de manière à ce que la totalité des phonèmes de la langue soient représentés. Pour cette étude pilote, une soprano professionnelle a dit le texte normalement (VPN = Voix Parlée Normale), puis le plus lentement possible (VPR = Voix Parlée Ralentie), et enfin le plus rapidement possible (VPA = Voix Parlée Accélérée). On lui a demandé ensuite de chanter l'extrait correspondant, à la cadence métronométrique, d'abord au tempo mentionné par le compositeur $\text{♩} = 80$ (VCN = Voix Chantée Normale) puis deux fois plus lentement $\text{♩} = 40$ (VCR = Voix chantée Ralentie) et enfin deux fois plus vite $\text{♩} = 160$ (VCA = Voix Chantée Accélérée). On lui a demandé par ailleurs de dire normalement une séquence de 24 syllabes puis de la chanter sur $\text{la}\#_2$ (233 Hz) pour le grave, $\text{la}\#_3$ (466 Hz) pour le médium et $\text{la}\#_4$ (932 Hz) pour l'aigu, en conservant le même débit qu'en voix parlée. Tous les enregistrements ont été effectués en chambre anéchoïque et ont fait l'objet d'une transcription phonétique large ainsi que d'une analyse acoustique de type ATFoCi (amplitude, temps, fréquence fondamentale, courbe d'intensité).

3. DUREE DES PHONEMES DANS LES PRODUCTIONS EXTREMES

3.1. Hypothèses

La langue française est caractérisée par son isochronisme syllabique, loi phonétique selon laquelle, les syllabes parlées tendent à être de durées égales. Dans le chant, l'isochronisme ne peut être maintenu. En effet, la durée moyenne des voyelles augmente d'une façon très importante (461% dans le corpus analysé) alors que la durée moyenne des consonnes diminue légèrement par rapport à la parole (7%) (Scotto

Di Carlo, 2005). Ceci se retrouve, toutes proportions gardées, en voix criée où la durée vocalique est augmentée de 50% et la durée consonantique réduite de 20% (Rostolland, 1979). Certains auteurs attribuent ce phénomène à l'effort vocal (Rostolland, 1982; Bonnot et Chevrier-Muller, 1991; Fónagy et Fónagy, 1966).

Il est toutefois difficile de se référer à la notion d'effort vocal car non seulement, elle diffère selon les auteurs mais le Sound Pressure Level (SPL) qui est traditionnellement utilisé pour le mesurer a été remis en question récemment par des chercheurs suédois qui ont démontré que le SPL ne joue pas un rôle majeur dans la perception de l'effort vocal (Rundlöf, 1996; Traunmüller, 1997; Eriksson et Traunmüller, 1999), ce qui les a amené à mettre au point une méthode de mesure plus fiable de l'effort vocal qui prend en compte l'aspect perceptif de celui-ci en intégrant un facteur proxémique (Traunmüller et Eriksson (2000). Kallopoulos (2000), quant à lui, a émis l'hypothèse selon laquelle dans le chant, l'augmentation de la durée des voyelles et la diminution de la durée des consonnes seraient liées à l'importance de la force articulatoire. Or, d'après une étude réalisée par Scotto di Carlo, (1979) c'est le phénomène inverse qui se produit lorsque la force articulatoire augmente, puisqu'on observe pour la voix chantée surarticulée, une légère diminution de la durée vocalique (14%) et une forte augmentation de la durée consonantique (58%). En outre, en raison des contraintes d'ouverture et d'aperture, la force articulatoire est moins importante dans le chant que dans la parole ainsi qu'en témoignent les études palatographiques réalisées sur des chanteurs professionnels, où du grave au médium, la langue laisse un contact de plus en plus léger sur le palais, et ne fait plus que l'effleurer dans l'aigu (Scotto di Carlo, 1994).

A ces hypothèses qui attribuent à l'effort vocal ou à la force articulatoire, le déséquilibre syllabique caractéristique des productions vocales réalisées dans des conditions extrêmes, il convient d'ajouter une observation qui ne concerne que la voix chantée et fait intervenir la notion de contrainte esthétique. En effet, dans le chant, l'abrègement des consonnes est davantage marqué du fait que si les voyelles constituent un support mélodique idéal, les consonnes interrompent le continuum sonore, ce qui va amener les chanteurs à réduire leur durée de manière à ce qu'elles nuisent le moins possible à la qualité du *legato*¹ (Scotto di Carlo, 1979). Cet abrègement consonantique varie également en fonction de la stratégie articulatoire du chanteur. Il est plus important pour les

voix aiguës qui évoluent dans une zone de fréquences où il devient difficile d'articuler (Scotto di Carlo, 1994).

3.2. Théories

Dans une étude comparative entre la voix parlée et la voix criée, Schulman (1989) rappelle les interprétations qui ont été données à l'accroissement des durées vocaliques.

- *Théorie mécaniste* : la vitesse des organes articulatoires ne pouvant augmenter proportionnellement à l'amplitude du déplacement, les organes ont besoin de davantage de temps pour effectuer leurs mouvements. (Lindblom, 1967)

- *Théorie communicationniste* : les voyelles sont les sons du langage les plus audibles à grande distance y compris dans un environnement très bruyant. Dans la mesure où ce sont elles qui véhiculent le plus d'information, elles sont intentionnellement allongées lorsqu'on crie. (Fónagy et Fónagy, 1966, Clark *et al.*, 1987).

La théorie mécaniste de Lindblom est la plus appropriée pour expliquer les phénomènes observés en voix chantée où les organes effectuent des déplacements beaucoup plus importants qu'en voix parlée et qu'en voix criée, ce qui nécessite davantage de temps pour les réaliser, ainsi que le prouve la séquence parlée et chantée avec le même débit syllabique pour laquelle dans le chant, l'augmentation de la durée totale par rapport à la parole est de 6.3% pour le grave, 13% pour le médium et 28,7% pour l'aigu.

Si, dans les conditions de production extrêmes, un allongement considérable des voyelles est observé par tous les auteurs, la durée des consonnes fait l'objet de controverses, probablement parce qu'il est difficile de faire la part entre les phénomènes dus à l'effort vocal et ceux dus à la force articulatoire, qui peuvent être dissociés comme c'est le cas pour le chant et qui ne produisent pas les mêmes effets temporels sur les phonèmes (Scotto Di Carlo, 1979)

4. EFFETS DU DEBIT ET DU TEMPO SUR L'EQUILIBRE SYLLABIQUE

La durée moyenne des consonnes comme celle des voyelles augmente lorsque le débit ou le tempo ralentissent, mais alors que les durées vocaliques augmentent dans des proportions importantes, les durées consonantiques varient dans une limite réduite.

Les modifications du débit et du tempo entraînent des changements dans la durée syllabique qui vont avoir pour effet d'altérer l'équilibre consonne/voyelle à l'intérieur de la

syllabe. Ainsi, plus le débit et le tempo ralentissent, plus la durée syllabique augmente et plus la proportion de consonne diminue.

4.1. Elasticité des voyelles

L'étude contrôlée du débit montre qu'en raison de leur grande élasticité, les voyelles parlées subissent des variations temporelles toujours plus importantes que celles des consonnes. On relève une exception cependant, pour la voix parlée accélérée (VPA) où les consonnes sont abrégées de 41% par rapport à la voix parlée normale (VPN) alors que les voyelles ne le sont que de 37%. L'une des explications possibles à cet abrègement plus important des consonnes est qu'à ce débit extrêmement rapide, elles sont nécessairement sous-articulées. Le débit du soprano est en effet de 6,65 syllabes/seconde en VPA contre 3,89 syllabes/seconde en VPN.

Pour la voix chantée, comme pour la voix parlée, les variations temporelles des voyelles sont toujours plus importantes que celles des consonnes. En effet, le taux d'expansion est de 76% pour les voyelles et de 38% pour les consonnes. Quant au taux de compression, il est seulement de 41% pour les voyelles et de 23% pour les consonnes. La théorie mécaniste permet de rendre compte de la difficulté éprouvée par le soprano en voix chantée accélérée, où son débit syllabique était seulement de 3,26 syllabes/seconde. La chanteuse n'a en effet pas réussi à accélérer suffisamment le tempo pour atteindre les 4,12 syllabes/sec escomptées, ce qui correspond dans ce cas particulier, à une compression de la durée syllabique de 80% seulement.

4.2. Elasticité des consonnes

Si les voyelles s'adaptent aux variations de débit et de tempo en se comprimant ou en se dilatant aisément, les limites de compression et d'expansion des consonnes sont restreintes en raison des fortes contraintes physiologiques qu'elles subissent. Les consonnes ont une élasticité plus ou moins importante selon leur nature. Dans la parole comme dans le chant, les occlusives sonores [b], [d] [g] qui subissent les contraintes aérodynamiques les plus fortes ont l'élasticité la plus faible (de 39 à 135 ms pour la parole et de 39 à 78 ms pour le chant). Cette valeur de 39 ms commune à la voix parlée et à la voix chantée, semble être la limite de compression des occlusives sonores au-dessous de laquelle on ne peut descendre, même en voix chantée accélérée où la valeur la plus brève relevée est égale à 38,85 ms. Quant aux consonnes vocaliques [l], [r], [R] dont les

contraintes sont très faibles, elles présentent l'élasticité la plus forte (de 34 à 224 ms en voix parlée et de 18 à 322 ms en voix chantée), ce qui en fait des phonèmes privilégiés dans le chant.

Les consonnes pour lesquelles l'élasticité est la plus importante sont généralement les plus faciles à chanter, car elles ne sont pas soumises à des contraintes physiologiques trop lourdes. C'est le cas par exemple des liquides [l] et [r] ainsi que des nasales [m] et [n] dont la structure acoustique très proche de celle des voyelles, facilite la cantillation. C'est d'ailleurs pour cette raison qu'on les utilise spontanément lorsqu'on chante un air dont on ne connaît pas les paroles.

5. LIMITES D'ELASTICITE DES CONSONNES ET DES VOYELLES

Les limites d'élasticité des consonnes sont pratiquement identiques dans la parole et dans le chant, la limite de compression étant de 16 ms pour la parole et le chant et la limite d'expansion passant de 324 ms pour la voix parlée à 340 ms pour la voix chantée.

5.1. Expansion et compression des consonnes

L'expansion de certaines consonnes comme les occlusives sonores par exemple, est limitée par des contraintes physiologiques de type aérodynamique. En effet, la fermeture complète du tractus vocal, qui caractérise la production de ce type de consonnes a pour effet d'accroître momentanément la pression supraglottique qui atteint sa valeur culminante au cours de la tenue consonantique. En raison des faibles variations de la pression subglottique, l'augmentation de la pression supraglottique entraîne une égalisation des pressions au-dessus et au-dessous de la glotte. Le débit d'air transglottique décroît alors brusquement et tend à s'annuler. La décroissance du débit transglottique provoque une diminution de l'effet de succion (effet Bernoulli) qui est le moteur principal du régime de vibration des cordes vocales. Par conséquent, plus on prolonge la tenue d'une occlusive sonore, plus la pression supraglottique augmente, plus le débit d'air transglottique décroît et plus l'effet Bernoulli est entravé, ce qui entraîne l'arrêt des vibrations laryngées et aboutit à leur dévoisement.

Les limites de compression des consonnes sont vraisemblablement liées à des contraintes physiologiques dues à l'inertie des organes articulatoires et/ou à leur mode de production comme c'est le cas par exemple, pour les

fricatives sourdes qui sont réalisées avec une glotte largement ouverte afin que les cordes vocales ne puissent entrer en vibration et un rétrécissement important du tractus vocal qui provoque des turbulences du flux aérien, responsables des bruits de friction qui les caractérisent. Or ce régime de turbulence met un certain temps à s'établir, d'une part parce que la pression supraglottique qui est très faible à cause de l'ouverture de la glotte ne peut augmenter qu'au moment où le tractus vocal est partiellement fermé par la constriction et d'autre part, parce que l'ajustement de cette constriction et du débit d'air transglottique doit être extrêmement précis afin que le flux aérien qui était laminaire puisse devenir turbulent en aval de la constriction. On comprend donc que les fricatives sourdes ne puissent être comprimées sans perdre leur identité.

5.2. Expansion et compression des voyelles

L'expansion des voyelles est très importante dans le chant où physiologiquement elle n'est limitée que par la tenue de souffle du chanteur et musicalement par la partition, à l'exception toutefois des voyelles nasales. En effet, lorsqu'on chante, on ne peut tenir une voyelle nasale car en raison de son inertie, le voile du palais s'affaisse sur le dos de la langue au bout de quelques centièmes de seconde, obstruant ainsi la cavité buccale. Pour éviter que la voyelle nasale devienne un son émis bouche exclue dont le timbre hypernasalisé est particulièrement inesthétique, les chanteurs utilisent un stratagème qui consiste à tenir la voyelle orale correspondante et à ne la nasaliser qu'au dernier moment.

Les limites de compression des voyelles sont également très proches en voix parlée et en voix chantée (respectivement 38 et 40 ms). Les raisons de ces limites de compressibilité sont d'ordre perceptif, une compression excessive des phonèmes étant incompatible avec leur intelligibilité.

6. CONCLUSION

Dans le chant, l'allongement syllabique très important entraîne une modification de la durée relative des consonnes et des voyelles à l'intérieur de la syllabe qui a pour effet d'augmenter le pourcentage des voyelles au détriment de celui des consonnes. La durée moyenne des consonnes comme celle des voyelles s'accroît lorsque le débit et le tempo ralentissent; mais alors que les durées vocaliques augmentent dans des proportions importantes, la durée des consonnes varie dans une limite réduite. Enfin, en raison des fortes contraintes

physiologiques auxquelles elles sont soumises, les consonnes chantées ont une durée intrinsèque et une élasticité très proches de celles des consonnes parlées.

En ce qui concerne les deux critères temporels qui caractérisent la voix chantée, l'allongement des voyelles est dû à un important ralentissement du tempo inhérent à toute production vocale, ainsi qu'à une augmentation de la durée articuloire liée à l'amplitude considérable des mouvements effectués par les organes phonatoires. Quant à l'abrègement des consonnes, il résulte des puissantes contraintes aérodynamiques qu'elles subissent au niveau de leur production mais aussi des contraintes esthétiques que s'imposent les chanteurs afin de préserver le *legato* qui est l'un des facteurs essentiels de la musicalité.

7. BIBLIOGRAPHIE

- [1] Bonnot, J.F., Chevrie-Muller, C. 1991. Some Effects of Shouted and Whispered Conditions on Temporal Organization. *J. Phonetics* 19 : 473-483.
- [2] Clark, J., Lubker, J., Hunnicutt, S. 1987. Some Preliminary Evidence for Phonetic Adjustment Strategies in Communication Difficulty. In Steele, R., Threadgold, T. (eds), *Essay in Honour of Michael Halliday*, Amsterdam : Benjamins.
- [3] Eriksson, A., Traunmüller, H. 1999. Perception of Vocal Effort and Speaker Distance on the basis of Vowel Utterances. *Proc. 14th ICPHS*. San Francisco, 2469-2472.
- [4] Fonagy, I., Fonagy, J. 1966. Sound Pressure Level and Duration. *Phonetica* 15 : 14-21.
- [5] Kallopoulos, N. 2000 *Syllable Temporal Patterns*. Doctoral Dissertation. National and Capodistrian University of Athens.
- [6] Lindblom, B. 1967. Vowel Duration and a Model of Lip Mandible Coordination. *STL-QPSR*, 4 : 1-29.
- [7] Rostolland, D. 1979. *L'audition de la parole en présence de bruit*. Thèse de Doctorat d'Etat. Paris.
- [8] Rostolland, D. 1982. Acoustic Features of Shouted Voice. *Acustica* 50 : 118-125.
- [9] Rundlöf, J. 1996. *Perceptuella letrådar vid auditiv bedömning av avståndet mellan talare och lyssnare*. D-Uppsats, Inst. för lingvistik. Stockholm universitet.
- [10] Schulman, R. 1989. Articulatory Dynamics of Loud and Normal Speech. *J. Acoust. Soc. Am.* 85, 1 : 295-312.
- [11] Scotto Di Carlo, N. 1979. Perturbing effects of Overarticulation in Singing. *J. Res. Singing* II, 2 : 10-27.
- [12] Scotto Di Carlo, N. 1994. L'intelligibilité de la voix chantée. *Médecine des Arts* 10 : 2-15.
- [13] Scotto Di Carlo, N. 2005. Structure temporelle de la syllabe parlée et chantée. *Médecine des Arts* 52 : 2-11.
- [14] Traunmüller, H. 1997. Perception of Speaker Sex, Age and Vocal Effort. *Phonum* 4 : 183-186.
- [15] Traunmüller, H., Eriksson, A. 2000. Acoustic Effects of Variation in Vocal Effort by Men, Women and Children. *J. Acoust. Soc. Am.* 107, 6 : 3438 - 3451.