

Caractéristiques temporelles de la parole imitée

Jana Mejvaldová

Institut de Phonétique, Université Charles
nám. J. Palacha 2, 116 38 Prague, République Tchèque
Tél.: + 420 221 619 250 ; mejvaldo@ff.cuni.cz

ABSTRACT

By comparing of the selected stretches of imitated speech with the imitation and the imitator's own voice, we tried to examine what acoustic characteristics may provide for successful imitation and, also, which prosodic qualities are stable and characteristic for a given speaker. In this study, we are focusing on comparing the overall duration and temporal proportions of the utterance segments examined. It appears that temporal correlation between the imitation and the utterance realized in the imitator's own voice is in most cases greater than that between the imitation and the original utterance. Overall duration also suggests the stability of temporal properties of the utterance for one speaker. F0 is more easy to imitate; the imitator may choose one of two strategies - to adjust the global pitch to the imitated utterance, or to imitate as precisely as possible the intonation patterns, while retaining the average F0 of his or her own natural voice.

1. INTRODUCTION

L'imitation de la parole est une discipline très populaire du divertissement radiotélévisé et des spectacles. Les études phonétiques dans ce domaine se montrent plutôt rares ; il n'en existe pas, à notre connaissance, pour le tchèque. Les recherches sur la possibilité d'imitation de la voix humaine par l'homme (son imitation par la machine entre dans un domaine que nous n'allons pas aborder) sont interdépendantes avec celles qui étudient les caractéristiques acoustiques du locuteur ([4], [5]). Leurs résultats peuvent être appliqués dans le ressort de la phonétique légale. En quoi consiste une imitation réussie? Quels sont les paramètres acoustiques qui caractérisent le locuteur? Et lesquels parmi eux sont pertinents pour son identification? Quand nous amusons-nous par une imitation réussie, et où est la limite au-delà de laquelle nous ne sommes plus capables de distinguer un discours original d'une imitation? Est-il d'ailleurs possible de répondre à ces questions par les recherches des qualités acoustiques de la parole?

2. BUTS DE L'ETUDE

Selon les recherches effectuées par les phonéticiens suédois Wretling, Eriksson ([3], [6]) et par Zetterholm ([7]) la distribution temporelle au niveau des segments de la parole se montre comme un phénomène difficilement imitable. Le niveau de F0, les variations intonatives et la durée globale du discours peuvent être imités par les imitateurs professionnels avec un succès remarquable. Il résulte des études citées ci-dessus que

la distribution temporelle au niveau des segments est peu variable dans le cadre d'un locuteur, bien qu'il s'agisse du locuteur-imitateur professionnel imitant une voix. Ces résultats, à condition que leur validité soit prouvée également pour le tchèque, seraient d'une grande valeur notamment pour la phonétique légale. Les études dans ce domaine sont susceptibles d'indiquer les phénomènes prosodiques stables pour un locuteur qui, ensemble avec les indices acoustiques conditionnés par les qualités physiologiques, le caractérisent.

Dans cette communication (qui fait partie d'une étude plus étendue), nous voulons aborder un problème partiel mais nullement marginal, qui est celui de la distribution temporelle dans la parole imitée : la relation temporelle de la parole imitée avec la voix d'origine et avec la propre voix d'imitateur.

Cette étude est basée sur l'analyse temporelle des performances imitatives d'un locuteur – imitateur tchèque célèbre. Son registre contient 260 voix. Nous allons nous concentrer sur les stratégies que l'imitateur adopte pour atteindre le modèle sonore le plus étroitement possible (tel a été le conseil à l'imitateur, et non pas de distraire les spectateurs), notamment dans le domaine de l'organisation temporelle du discours. Nous mentionnerons brièvement le comportement de la fréquence fondamentale dans la parole imitée et constaterons les rapports de ce paramètre entre la parole imitée et l'imitation et entre l'imitation et la propre voix d'imitateur.

3. MATERIEL

Pour ce type d'observation, nous disposons du matériel sonore suivant : enregistrement de deux voix d'origine, destinées dans notre étude à être imitées, selon le choix de l'imitateur. Les enregistrements des voix d'origine proviennent des archives de la Radio tchèque. Les enregistrements ont été transcrits avec l'intention de procurer à l'imitateur un support écrit pour ses performances. Nous avons noté des sons d'hésitation, des lapsus etc. Les textes transcrits, ensemble avec les enregistrements sonores, ont été fournis à l'imitateur afin qu'il puisse se familiariser avec le texte avant d'effectuer l'enregistrement expérimental.

Les tests auditifs pilotes ont montré que le support écrit qui a dû conduire le locuteur dans ses performances lui a fait plutôt obstacle ; les résultats sont plus persuasifs lorsque l'imitateur est libre dans

sa production, quand il peut improviser et quand le contenu lexical et la structure syntaxique ne sont pas définis d'avance. Cette étude est néanmoins ciblée sur les performances imitatives réalisées avec un support écrit. Nous aurons ainsi la possibilité de comparer les proportions temporelles du discours d'origine, de l'imitation et de la voix propre de l'imitateur. Le Tableau 1 résume les caractéristiques acoustiques élémentaires des discours de deux locuteurs imités, de leurs imitations et du discours émis par la voix de l'imitateur. Il s'agit évidemment du même texte. Le nombre plus élevé de syllabes dans la parole imitée est dû au fait que l'imitateur a répété (en imitant le locuteur L1) ou inséré (dans le cas du locuteur L2) certains mots. La répétition des mots est caractéristique pour le locuteur L1 mais elle ne paraît pas dans le discours observé. Ce fait a contribué à l'hypothèse que l'imitateur compte sur l'expérience commune et sur la connaissance du locuteur imité partagée par l'imitateur et par les auditeurs. Dans ses imitations, l'imitateur fait souvent l'allusion aux phénomènes caractéristiques pour la parole du locuteur imité, sans qu'ils soient présents dans le discours immédiatement imité et analysé. Il ne s'agit pas obligatoirement des phénomènes acoustiques.

L1	<i>voix d'origine</i>	<i>parole imitée</i>	<i>voix propre de l'imitateur</i>
Durée totale	36,93	50,4	35,94
nombre de syllabes	146	166	147
débit articulatoire/ débit de parole	4,4/3,95	4,06/3,3	4,6/4,09
Moyenne de F0	105,4 Hz	120,9 Hz	119,73 Hz
L2			
Durée totale	46,49	55,42	48,67
nombre de syllabes	172	176	189
débit articulatoire/ débit de parole	4,19/3,69	3,7/3,18	4,2/3,88
Moyenne de F0	143 Hz	147 Hz	144 Hz

Tableau 1 : Caractéristiques acoustiques des discours entiers de deux locuteurs imités.

3.1. Moyens techniques

Les enregistrements des imitations et de la voix propre de l'imitateur ont été acquis dans la chambre anéchoïque de l'Institut de Phonétique à Prague, enregistrés sur une cassette DAT et numérisés avec le logiciel Sound Forge. L'étiquetage du signal, les mesures temporelles et fréquentielles ont été effectuées avec le logiciel Praat 4.0.49.

4. ANALYSE TEMPORELLE

Pour l'analyse détaillée des proportions temporelles (et de la courbe de F0 dans la deuxième partie de l'article) nous avons choisi comme exemples cinq unités prosodiquement et sémantiquement complexes (3 exemples de L1 et 2 exemples de L2). Les unités ont été segmentées au niveau des mots, les mots ajoutés pendant l'imitation ont été enlevés.

4.1 Durée totale

La durée totale et proportionnelle des exemples choisis pour l'analyse détaillée est présentée dans la figure suivante (Figure 1). Chaque triade représente une unité analysée : trois premières triades du haut renvoient au locuteur L1, les deux suivantes au locuteur L2. La durée des segments se lit en abscisse. Dans chaque triade, la première bande est réservée pour *la voix imitée* et ses proportions temporelles. En deuxième bande de chaque triade, la durée totale avec les proportions des segments de *l'imitation*, suivie de la représentation temporelle du même discours réalisé par *la voix propre* de l'imitateur (troisième bande). Les informations concernant la durée globale de la parole imitée et de ses proportions temporelles qui sont schématisées dans la Figure 1 témoignent un phénomène persistant : avec deux exceptions (première triade, L1/1 et dernière triade, L2/2), la durée totale de l'imitation est systématiquement plus proche de la voix propre de l'imitateur que de la parole imitée. Dans trois cas, la durée totale de l'imitation excède celle de la voix imitée. Nous proposons l'hypothèse que ce rallongement ou plutôt ralentissement est causé par l'effort articulatoire et phonatoire élevé.

4.2 Corrélation de la durée des segments

Le coefficient de corrélation entre la durée des segments dans deux discours indique le degré de leur similarité. Dans les phrases que nous avons analysées, la corrélation temporelle entre la voix imitée et l'imitation a été calculée ainsi que la corrélation entre l'imitation et la voix propre de l'imitateur. Les corrélations sont présentées dans le tableau 2.

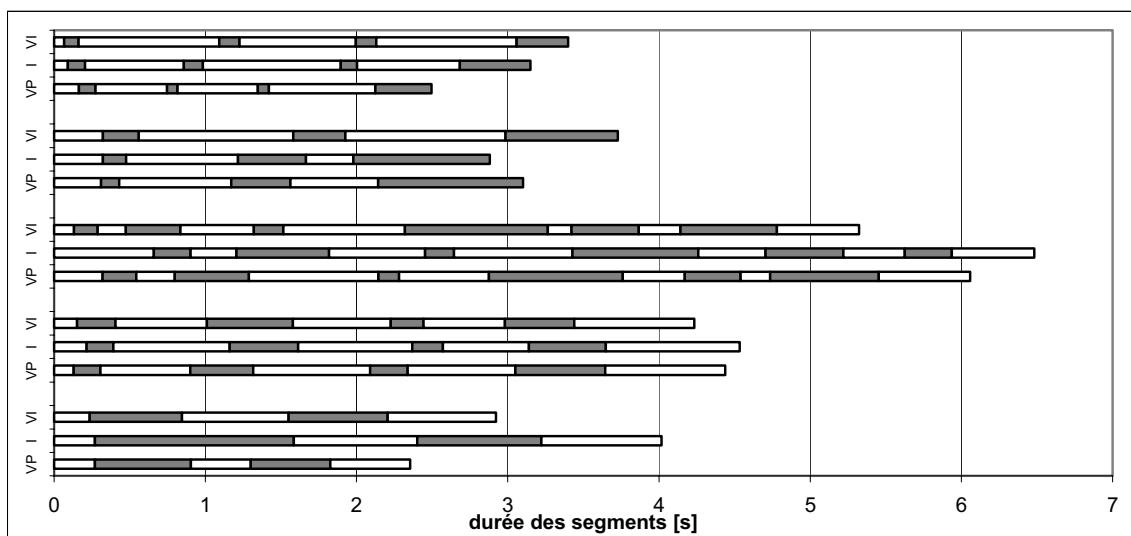


Figure 1 : Proportions temporelles des exemples analysées. En haut de chaque triade la voix imitée (VI), suivie de l'imitation (I) et de la voix propre de l'imitateur (VP). Les triades correspondent aux exemples choisis pour l'analyse détaillée. Chaque champs représente un segment (mot) de la phrase analysée. L'alternance des champs blancs et gris sert à une meilleure orientation dans la figure.

	<i>Corrélation voix imitée/imitation</i>	<i>Corrélation imitation/voix propre</i>
L1/1	0,62	0,71
L1/2	0,80	0,84
L1/3	0,64	0,66
L2/1	0,95	0,93
L2/2	0,68	0,92

Tableau 2 : Corrélation entre les voix imitées des locuteurs L1 et L2 dans les unités analysées (L1/1 – L2/2) et les imitations et entre les imitations et la voix propre de l'imitateur.

Avec une exception (en gris), dans tous les cas la corrélation entre la durée des mots dans le discours prononcé par la voix propre de l'imitateur et en imitation est plus élevée que la corrélation entre l'imitation et la voix imitée. Pour comparaison – une de nos études précédentes concernant l'acquisition de la prosodie d'une langue étrangère par imitation ([2]) indique que la corrélation avec la parole imitée dans une langue étrangère pour un locuteur non entraîné tourne autour de 0.3.

5. FREQUENCE FONDAMENTALE

Selon les résultats des études effectuées par les chercheurs suédois ([3], [6], [7]), le comportement de la fréquence fondamentale et de ses valeurs moyennes s'avèrent comme relativement facilement imitables. Nous avons vérifié, sur le matériel des mêmes unités analysées, le fonctionnement de l'imitation dans le domaine de F0.

La Figure 2 présente les courbes mélodiques de deux unités analysées, chacune provenant d'un locuteur imité. Le discours imité est en haut, suivi du schéma de l'imitation et de l'intonation de la voix propre de l'imitateur en dernier lieu. La comparaison entre la voix imitée et la voix propre de l'imitateur suggère que la conception syntaxico – prosodique de deux locuteurs a été différente. Dans l'imitation l'imitateur s'approche de la courbe intonative imitée. La moyenne de F0 dans les deux exemples et ses changements selon l'intention de l'imitateur sont résumés dans le Tableau 3. Dans le premier exemple, l'imitateur a mieux réussi de s'approcher du niveau global de F0, alors que dans l'exemple second il a préféré garder approximativement le niveau fréquentiel de sa voix naturelle ; en revanche, il a mieux adopté la courbe intonative du discours imité.

Exemple	<i>Voix imitée</i>	<i>Imitation</i>	<i>Voix de l'imitateur</i>
L1/2	102 Hz	109 Hz	122 Hz
L2/2	175 Hz	150 Hz	154 Hz

Tableau 3 : F0 moyen dans la parole imitée, dans l'imitation et dans la voix propre de l'imitateur.

6. CONCLUSION

Les résultats de ce sondage dans le domaine des caractéristiques acoustiques et surtout temporelles de la parole imitée sont en accord avec les informations acquises par les recherches réalisées pour le suédois (Wretling et Eriksson [6], Zetterholm [7]). Il en résulte que les propriétés temporelles, la durée globale ainsi que la durée proportionnelle des segments, s'avèrent comme les caractéristiques stables du locuteur. Ce sont des phénomènes difficilement imitables et les imitateurs soit restent dans leur habitudes discursives, soit exagèrent jusqu'à la caricature. Dans la plupart des cas, la corrélation temporelle entre l'imitation et le discours réalisé par

la propre voix d'imitateur est plus élevée que celle entre l'imitation et la parole imitée. Le comportement de F0 est imitable plus facilement. L'imitateur peut adopter deux stratégies : approcher le niveau global de sa voix vers le F0 moyen de la parole imitée ou bien suivre le plus exactement possible les contours intonatifs tout en gardant le niveau de F0 de sa voix naturelle. Ce type d'étude exige les analyses d'une quantité énorme pour que les résultats puissent être considérés comme solides et fiables. Le sondage dont les résultats viennent d'être présentés fait une partie préliminaire d'une étude plus vaste de longue durée.

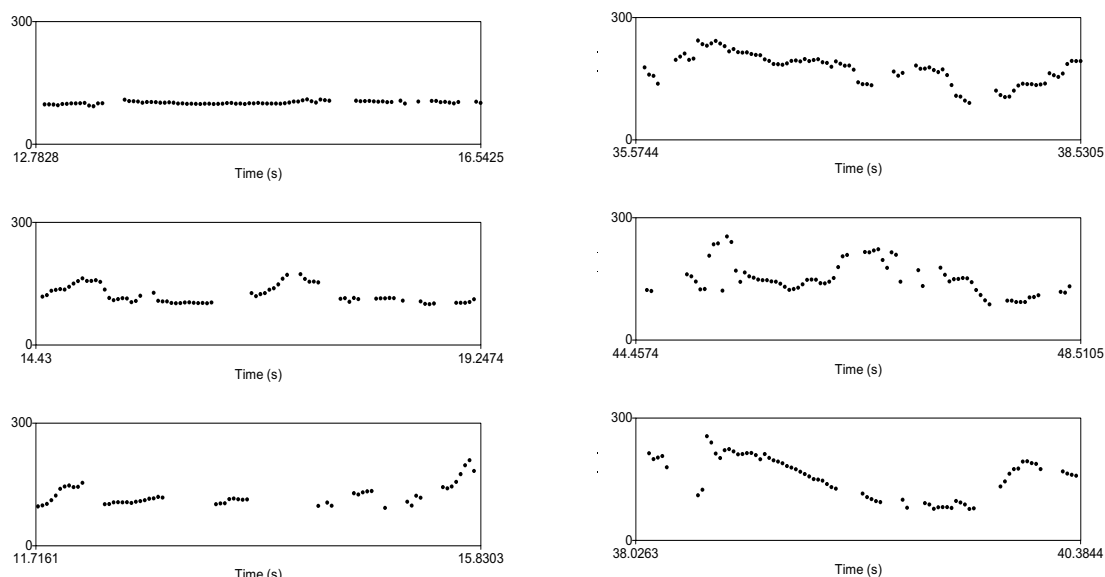


Figure 2 : Courbes de F0 dans deux phrases analysées. A gauche l'exemple L1/2, à droite l'exemple L2/2. Du haut en bas : parole imitée, imitation, enregistrement de la voix propre de l'imitateur.

REFERENCES BIBLIOGRAPHIQUES:

- [1] P. Bessler. La caricature de de Gaulle par Tissot: Etude phonostylistique. In: *Information/Communication*, 12, Candian Scholars' Press, pp. 19 – 32, 1991.
- [2] M. Dohalská – T. Duběda – J. Mejvaldová. L'imitation rythmique et ses limites: analyse de la structure temporelle, *Colloque de la Société Internationale de la Linguistique Fonctionnelle*, České Budějovice (sous presse), 2003.
- [3] A. Eriksson. – P. Wretling. How flexible is human voice? – Acase study of mimicry, In: *Proceeding of Eurospeech '97, Rhodos*, vol 2, pp. 1043 – 1046, 1997.
- [4] P. Janota. *Personal Characteristics of Speech*, Rozpravy ČSAV, 77, 1, Academia, Praha, 1967.
- [5] M. Svobodová. Individuální variabilita vybraných řečových parametrů v řeči českého mluvčího, doktorská dizertační práce, nepublikováno [Variabilité individuelle des paramètres choisis dans la parole du locuteur tchèque, thèse de doctorat, non publié], 2000.
- [6] P. Wretling – A. Erikson. Is articulatory timing speaker specified? – Evidence from imitated voices, tiré-à-part.
- [7] E. Zetterholm. Intonation Patterns and Duration Differences in Imitated Speech, In: Bel, B. & Marlien, I. (Eds.) *Proceedings of Speech Prosody*, pp. 731 – 734, ISBN 2-9518233-0-4, 2002.
- [8] E. Zetterholm. An Impersonator and some of his voices. A phonetic study of human voice flexibility, tiré-à-part.
- [9] E. Zetterholm. The same but different – three impersonators imitate the same target voice. In: *Proceedings of the 15th ICPHS, Barcelona*, pp. 2205 – 2208, ISBN 1-876346-48-5, 2003.