

Styles de prononciation et types de diction

Lija BONDARKO, Ljudmila VERBICKAJA, Mirra GORDINA,
Vadim KASEVIČ, Lev ZINDER

La notion de style de prononciation occupe réellement une place périphérique dans la phonétique, bien que plusieurs auteurs abordent cette question dans une certaine mesure (Avanesov 1972 ; Léon 1971 ; Šapiro 1968)¹. La distinction des styles de diction est considérée comme importante pour la théorie et la pratique de l'enseignement de langues, et pourtant, même concernant cet objectif, dans la majorité des cas, il n'existe pas de recommandations élaborées.

On estime que l'opposition entre un style «plein» [*polnyj*] et un style «parlé» [*razgovornyj*], établie par Lev Ščerba en 1915 (Ščerba 1957 : 21)², se situe au même rang que la distinction des registres solennel, oratoire, officiel, familial, etc. La classification proposée par Ščerba est considérée seulement comme plus généralisée.

La raison en est donnée par Ščerba lui-même. Premièrement, en utilisant le terme de «style», ce-dernier a provoqué une interprétation de ce phénomène trop vaste, et même fausse. Deuxièmement, dans certaines de ses citations, il inclut ces styles dans les styles de parole en général, comme, par exemple, dans les propos suivants :

Les diverses formes de la parole, adaptées aux différentes conditions sociales, s'appellent «styles de parole». Les différents styles se distinguent les uns des autres non seulement lexicalement et syntaxiquement, mais aussi phonétiquement. Conformément à la grande diversité des conditions sociales, on peut distinguer quantité de styles. Par souci de simplicité, nous en distinguerons seulement deux. (Ščerba, 1937, p. 19)

¹ Dans le travail de Šapiro, on trouve une distinction entre «style» et «code», qui est proche de la distinction entre style de prononciation et type de diction.

² Dans cet article, on ne trouve pas encore les termes de «polnyj» [*plein, complet*] et «razgovornyj» [*parlé*], mais l'opposition elle-même est déjà clairement formulée.

Par ailleurs, on peut trouver chez Ščerba encore d'autres propos. Par exemple, dans un de ses derniers travaux, publié de façon posthume, nous trouvons une formulation qui se différencie significativement de celle citée ci-dessus :

Il est tout à fait évident qu'est ici possible un nombre infini de degrés intermédiaires, de la clarté et la netteté absolue (par exemple, par une prononciation syllabe par syllabe) jusqu'à la précipitation négligée, quand toutes les syllabes non-accentués sont à moitié avalées. En pratique, il est suffisant de distinguer deux types de prononciation. Nous appellerons le premier «style plein», l'autre «style parlé». (Ščerba 1957, p. 157)

Ici, comme nous le voyons, il ne s'agit pas uniquement de netteté et de non-netteté de la prononciation. Toutefois, dans la mesure où le terme de «style» persiste, existait et existe encore une possibilité de confusion des phénomènes, différents par leur nature.

Evidemment, il est raisonnable de distinguer d'un point de vue terminologique ces deux conceptions du «style de prononciation», en réservant ce dernier terme exclusivement à la caractéristique phonétique commune de l'énoncé. Pour renvoyer à la caractéristique phonétique de détail des éléments de la parole, et jusqu'à celle des segments de l'extension phonémique, on peut proposer le terme de «type de diction»³.

Combien existe-t-il de types de diction et comment se distinguent-ils ? Ščerba discernait deux styles de prononciation (dans nos termes – deux types de diction): plein et parlé, ou — plutôt — «incomplet». La dichotomie de cette classification possède un caractère de principe : elle découle du fait que ce n'est pas telle ou telle qualité d'exécution phonétique ou tel ou tel degré de perfection qui doivent jouer le rôle de critère distinctif des types de diction plein ou incomplet, mais bien plutôt la possibilité d'une *interprétation phonémique* [souligné par nous] d'un segment de la parole. On parlera de type de diction plein dans le cas où un segment de la parole reçoit une certaine interprétation phonémique. Si cette interprétation phonémique unique est impossible, nous avons alors devant nous un type incomplet.

Par exemple, le mot *časny* ['les heures'] peut être prononcé de telle façon que l'identité de la voyelle de la première syllabe comme allophone des phonèmes respectivement /i/, /e/ ou /a/ soit peu claire. En même temps, il ne fait aucun doute que ce mot peut être prononcé également de façon à ce que cette ambiguïté soit exclue. Cela peut être établi objectivement aussi bien par une analyse articulatoire-acoustique que, et c'est particulièrement important, par la perception de ce mot par les locuteurs natifs russes

³ Les termes de «prononciation» ['*proiznošenie*'] et de «diction» ['*proiznesenie*'] n'ont pas le même sens : le premier désigne un phénomène constant pour un individu ou un collectif donnés, alors que le deuxième caractérise un acte de parole, analysé uniquement du point de vue de sa face sonore.

[‘*nositeli jazyka*’]. Dans le premier cas, nous avons affaire à un type de diction incomplet, et dans le deuxième cas à un type plein.

Ainsi, le critère de définition du type plein est la reconnaissance univoque de la composition phonémique d’une séquence de la parole donnée, et celle-ci, à son tour, est assurée par le maintien dans la prononciation des caractéristiques phonétiques essentielles (pour ce dernier fait, Ščerba utilisait le terme de «composition phonétique idéale»). Le type plein de diction permet *toujours* [souligné par nous] de définir la composition phonémique d’un segment donné de la parole (et également *toujours* de répéter le segment entendu). Le type incomplet, au contraire, impose, pour cette tâche, un contexte ou une situation concrets.

Par conséquent, aucun énoncé ne peut appartenir *entièrement* [souligné par nous] au type incomplet (à l’exception, probablement, des répliques complètement déterminées par leur contexte ou par leur situation). Autrement dit, le type plein doit obligatoirement caractériser la diction d’au moins certaines parties de n’importe quel énoncé. Si cela n’était pas le cas, si tous les segments de l’énoncé permettaient simultanément une pluralité d’interprétations, la perception n’aurait pas de point d’appui permettant la compréhension du texte.

En même temps, même des observations superficielles de la parole coordonnée témoignent que de loin pas tous les segments de cette dernière se réalisent comme relevant du type plein de diction. Il convient, en rapport avec cela, de citer les propos de Ščerba tirés de son article «Des différents styles de prononciation et de la composition idéale phonétique du mot» :

La face sonore du mot, qui semblait toujours si claire, si indiscutable, et qui se présentait comme un noyau donné de représentations sémiologiques plus au moins floues, se trouve, donc, elle-même non moins floue et indéfinie. (Ščerba, 1957, p. 21)

Une idée analogue revient à R. Jakobson et à M. Halle, qui écrivent :

Puisque dans beaucoup de cas la charge fonctionnelle du phonème se trouve en réalité réduite pour l’interlocuteur, le sujet parlant, à son tour, est déchargé de la nécessité de respecter exactement toutes les différences sonores ; [...] L’image sonore de la parole peut être non moins elliptique que sa construction syntaxique. (Jakobson, Halle, 1962, p. 234)

Ainsi, la chaîne de la parole est habituellement hétérogène. Elle est formée par des segments du type plein de diction, qui s’associent dans la perception de façon univoque aussi bien avec certains phonèmes concrets (ou bien des séquences des phonèmes) qu’avec des segments de type incomplet. Du point de vue de la qualité objective des segments, le type incomplet de diction permet l’existence d’une large gradation. Une chute totale d’un de ces segments peut également se produire. L’indétermination phonétique survenant dans ces cas se caractérise par des niveaux divers de sévérité et peut être éliminée soit par l’extension du contexte, soit par la

limitation du cercle des solutions possibles (par exemple, par l'absence de mots avec une composition phonémique proche).

L'hétérogénéité objective de la chaîne de la parole (manifestée dans certaines limites) n'a pratiquement pas d'influence sur la perception de la parole. Non seulement la perception des segments du type plein, mais aussi la perception des segments qui se caractérisent par un type incomplet de diction restent en général catégorielles, et non fondées sur les lois de la logique des probabilités; s'établit en résultat une séquence de phonèmes tout à fait définie et qui correspond à une unité significative donnée. En plus, l'indétermination de la diction de la chaîne de la parole n'est même pas normalement fixée par la conscience. En conséquence, une telle chaîne de la parole doit être considérée comme homogène du point de vue fonctionnel.

L'indétermination objective de la chaîne de la parole se manifeste (si on l'examine du point de vue de la perception) à l'évidence non dans le résultat final de l'activité d'analyse de l'être humain, mais plutôt dans les méthodes de prise de décision de ce dernier. Les segments prononcés avec le type plein peuvent être reconnus grâce à leurs caractéristiques objectives. La décision concernant les segments du type incomplet est prise en résultat d'autres procédures, basées non seulement sur l'analyse de l'image phonétique, mais aussi (et parfois davantage) sur la prise en compte des liens notamment grammaticaux, sémantiques, et même extralinguistiques.

Ainsi, la perception de la parole, même au sens le plus étroit, c'est-à-dire en tant qu'identification de la composition phonémique de l'énoncé, ne peut pas être attribuée uniquement à l'aspect phonologique. Strictement parlant, ce dernier aspect s'occupe de la reconnaissance uniquement des faits appartenant au type plein de diction. Dans le cas du type incomplet, l'identification d'une unité verbale donnée non seulement ne résulte pas de l'identification de sa composition phonémique, mais, au contraire, souvent l'identification de la composition phonémique est seulement un produit dérivé de la reconnaissance d'une unité donnée grâce à la redondance contextuelle.

Ce qui vient d'être dit ne contredit pas l'opposition des types de diction. Il reste indiscutable qu'il doit exister un certain minimum de segments de la parole dont la composition phonémique ne nécessite pas pour son identification d'information complémentaire. Par exemple, dans une situation d'introduction d'un nouveau mot totalement inconnu des auditeurs, la distinction des types de diction se manifeste de façon tout à fait évidente: dans ce cas, il est absolument nécessaire d'utiliser le type plein dans la prononciation de l'élément radical et des affixes dérivationnels, tandis que la prononciation des affixes flexionnels prédits par le contexte ne le nécessite pas.

Normalement, en présence du contexte, la reconnaissance des segments du type plein peut également avoir lieu à l'aide de facteurs extra-phonétiques, de sorte que les analyses phonétique, grammaticale et sémantique d'un segment perçu de la parole se déroulent simultanément.

La réalité de l'existence de différentes procédures de perception des séquences phonétiques peut être considérée suffisamment évidente. En dépit de l'homogénéité des résultats de ces diverses procédures, n'est phonologique que la procédure de perception des segments du type plein, fondée sur l'usage d'une information proprement phonétique. Si, par exemple, un segment complètement absent dans la prononciation est restitué automatiquement grâce aux facteurs du sens, cela n'a aucun rapport ni avec les procédures phonologiques ni, par conséquent, avec la phonologie.

Ce ne sont pas uniquement un segment donné et son entourage direct qui portent une information phonétique pertinente. Le modèle prosodique de la syllabe, du mot, du syntagme, etc., joue également un rôle important. Pour le russe – en premier lieu, naturellement, pour les voyelles – la position du segment par rapport à l'accent est significative. Dans les expériences de L.V. Bondarko et de L.A. Verbickaja (Bondarko, Verbickaja 1973), il a été observé que, par exemple, la dernière voyelle du mot *няня* [няня 'nourrice', nom.sg.], en situation isolée, est perçue constamment par les auditeurs comme /i/. Cependant, dans la structure du mot, la même voyelle est perçue, de façon tout aussi constante, comme /a/ (et cela ne peut pas être expliqué par la domination de la forme de base, c'est-à-dire de la forme du dictionnaire, car le mot *няни* [няни 'nourrice', gén.sg.] a été perçu avec /i/).

Cependant, il serait prématuré de penser, sur la base de tout ce qui vient d'être présenté, que le critère d'une reconnaissabilité stable s'avère inefficace. Les données expérimentales décrites peuvent être expliquées de la façon suivante. Pour prendre une décision sur l'attribution phonémique d'une voyelle russe, l'individu doit utiliser l'information sur le modèle prosodique du mot; les règles correspondantes qualifient le son non seulement sur la base de la mesure de certains paramètres acoustiques, mais aussi sur la base d'un savoir sur la corrélation de ces paramètres avec le lieu du segment par rapport à l'accent. Dans le cas de l'isolation de la voyelle, comme c'était le cas dans l'expérience, l'information sur son caractère non-accentué est retirée, ce qui provoque une divergence de perception.

En conséquence, une perception univoque témoigne de la présence du type plein de diction seulement dans les cas où sont assurées les conditions d'utilisation d'une information phonétique pertinente. Avec cela, il faut définir comme information phonétique pertinente les caractéristiques de la chaîne de la parole, dont il est évident qu'elles sont utilisées par les locuteurs natifs pour l'identification phonémique. Ainsi, les règles de coarticulation et de réduction doivent-elles sans aucun doute être considérées comme cette information phonétique pertinente, dans la mesure où, si celles-ci ne sont pas prises en compte, ni la production, ni la reconnaissance de l'image phonétique du mot ne sauraient être assurées.

La perception univoque en elle-même ne signifie pas encore que nous avons obligatoirement affaire à un segment de parole caractérisé par

le type plein de la diction, car une telle perception est souvent assumée par l'usage d'une information extra-phonétique. On peut affirmer un tel lien seulement à condition que soit retirée la possibilité d'utilisation d'une information extra-phonétique, c'est-à-dire, de fait, quand l'énoncé est privé de sens (ou alors quand l'interprétation sémantique de l'énoncé est rendue possible seulement après identification de sa composition *compète* en phonèmes).

Du point de vue pratique, cela signifie que le matériel le plus fiable pour l'étude des corrélats phonétiques des unités phonologiques sont les «mots» phonologiques sans sens, c'est-à-dire, des «mots» privés de sens mais construits selon les règles phonotactiques de la langue donnée. On peut imaginer le schéma suivant de préparation du matériel expérimental: on compose des mots phonologiques (syntagmes, propositions) qui sont présentés pour leur production à un speaker. Pour s'assurer que ce speaker ait prononcé le matériel expérimental en utilisant le type plein, un enregistrement est présenté aux auditeurs; après cela on peut alors procéder à l'analyse de l'image phonétique objective, en étant assuré que dans ce matériel sont absents des segments dont l'attribution ne se soumet pas en réalité à une détermination exacte.

Tout cela ne signifie pas qu'il ne faut pas tendre vers l'éclaircissement de l'image réelle de la parole coordonnée spontanée. Il est simplement nécessaire de prendre en compte que dans celle-ci, un pourcentage considérable de la chaîne de la parole est représenté par des segments du type incomplet, dont la qualification d'un point de vue phonologique est irréalisable⁴. D'un côté, cela signifie que l'étude de tels segments ne donne pas d'information sur les caractéristiques invariantes des phonèmes dans le système; d'un autre côté, si nous possédons telle information, en recevant les données sur les paramètres des segments concrets du type incomplet de diction de la chaîne de la parole, nous ne pouvons pas mettre ces deux groupes des données en corrélation mutuelle univoque.

Bien évidemment, il est impossible de ne pas prendre en compte la grande variabilité des caractéristiques des phonèmes dans le système. De ce point de vue peut survenir la nécessité d'opérations relativement complexes sur la chaîne de la parole pour l'identification des phonèmes possédant des caractéristiques présentant un large «champ de dispersion». Autrement dit, l'absence de règles *simples* [souligné par nous] d'identification ne signifie pas l'impossibilité de principe de cette dernière. Cela conduit à l'idée de la nécessité d'opérations d'analyse complexes sur la chaîne de la parole pour la définition du «champ de dispersion» des caractéristiques des phonèmes.

Il devient évident que la notion de type de diction joue un rôle capital pour la méthode de détermination de la composition phonémique aussi bien de la langue en général que d'unités langagières données.

⁴ Le problème de la reconnaissance automatique de la parole est, comme en témoigne ce qui vient d'être exposé, beaucoup plus compliqué qu'il se présente d'habitude.

Pratiquement, en phonologie, cette notion est complètement exclue ; on considère qu'il faut s'appuyer sur la parole coordonnée simultanée, surtout dans la mise en évidence de la composition des phonèmes des mots et des morphèmes. Cependant, déjà en 1915, Ščerba reliait l'analyse des phonèmes avec «la composition phonétique idéale des mots». Jakobson et Halle quant à eux écrivent :

La diction négligente est seulement une variante de la production réduite de la forme complète et exacte de la parole, qui porte le plus grande quantité d'information. Dans l'analyse du système des phonèmes et des traits distinctifs qui les composent, il faut partir du code le plus complet qui existe dans la disposition des locuteurs parlant cette langue donnée. (Jakobson, Halle, 1962, p. 234-235)

De quelle façon peut-on se représenter la correspondance entre données phonologiques, obtenues sur la base de l'analyse d'un matériel de préférence artificiel, et données de la parole normale ? Nous partons de l'idée qu'une certaine partie des segments d'un énoncé doit être obligatoirement réalisée dans le type plein de diction. Ainsi pourra-t-on définir quel pourcentage de quels segments doit être prononcé de telle manière que l'énoncé puisse être interprété du point de vue phonologique, c'est-à-dire identifié avec les unités verbales concrètes.

Outre cela, il est raisonnable d'étudier le degré de déformation acceptable des segments du type incomplet de la diction dans certaines conditions données. Le degré d'une telle déformation peut fortement varier : d'un son phonétiquement intermédiaire par rapport au phonème paradigmatiquement le plus proche jusqu'à la chute complète du son. Il est évident que le pourcentage des segments de type plein et le degré de déformation des segments de type incomplet se trouve dans une certaine corrélation.

On peut de la même manière définir la caractéristique de la prononciation de certains locuteurs.

La variation de la part et de la qualité des segments du type plein et du type incomplet dans un texte est non seulement déterminée par l'exigence d'assurer l'intercompréhension, mais elle est aussi liée aux conditions de la communication, qui déterminent le *style* [souligné par nous] fonctionnel de la prononciation. Chaque style trouve sa manifestation également dans les caractéristiques phonétiques – particulièrement dans la part relative des segments du type plein et du type incomplet de la diction, ainsi que, peut-être, dans leur qualité spécifique et dans leur distribution dans le tissu syntaxique d'énoncé.

La plus ou moins grande révélation des caractéristiques données pourra justifier l'attribution de l'énoncé donné à tel ou tel style. Il serait ainsi possible d'établir une sorte d'échelle qui permettra de définir qu'un énoncé contenant un certain pourcentage de segments du type plein est un énoncé, par exemple, du style parlé, et qu'un énoncé avec un autre poids spécifique de segments de tel type doit être attribué à un autre style

fonctionnel «plus haut» ou «plus bas» du style parlé. Par cela même, la phonostylistique aurait à sa disposition un certain instrument pour la distinction objective des styles fonctionnels.

La distribution des segments des types plein et incomplet dans la chaîne de la parole, évidemment, est soumise à certaines règles. Sans toucher les tâches expressives et les autres tâches, qui jouent un rôle important dans le choix du type de diction, il faut, avant tout, indiquer le lien naturel entre les segments du type plein et la position forte : en règle générale, le type plein de diction marque, tout d'abord, la syllabe accentuée. Au contraire, les réductions qualitatives et quantitatives, qui caractérisent les positions faibles (non-accentuées) créent les conditions pour la réalisation du type incomplet. Cela, d'ailleurs, ne signifie pas qu'un son réduit obligatoirement soit un élément du type incomplet. Ainsi, dans la parole russe, une voyelle d'avant accentuée [i], malgré sa brièveté relative et une certaine réduction qualitative, peut garder une attribution phonémique unique.

Une étude plus complète des rapports entre les segments des différents types de diction dans la parole permettrait de mettre en évidence un certain «schéma du squelette» du mot, sur lequel se base la perception. On remarque ici la possibilité d'interpréter la notion connue d'image générale du mot du point de vue de son *intégralité* [souligné par nous]. Il est connu que l'intégralité phonétique du mot dans des langues avec l'accent est formée par l'accent même, qui exerce, avant tout, les fonctions cumulative et unificatrice. La mise en évidence d'une syllabe accentuée crée une *hiérarchie* de la structure phonétique du mot, qui à son tour assure une perception plus économe. Une perception strictement successive, «segment à segment», serait, évidemment, moins économe (même dans le cas d'un usage simultané d'une information extra-phonétique), qu'une sorte de perception par niveaux orientée vers certains points de repère. On peut penser que ces «points de repère» sont non seulement la syllabe accentuée, mais aussi la structure rythmique du mot entier, ainsi que la première syllabe du mot (cela explique l'assez régulière forte intensivité de la voyelle initiale). Sur le même plan, on peut analyser aussi l'inégalité informative des voyelles et des consonnes.

Pour préciser la notion de type de diction, il est important de prendre en compte, que le type plein dans lequel se réalise l'image phonétique idéale de l'unité langagière ne possède aucun lien avec la norme orthoépique. Le mot *sem'* ['sept'], par exemple, peut être prononcé de deux façons avec la préservation du type plein : avec une consonne dure ou avec une consonne molle en finale (*/s'em/* et */s'em'/*) ; le fait de déterminer quelle prononciation est correcte relève de la norme, de l'orthoépie. De même façon, le mot *pamjatnik* [памятник 'monument'] peut présenter différentes prononciations de la deuxième voyelle (*/pám'itn'ik/, /pám'etn'ik/, /pám'atn'ik/*) ; et peu importe laquelle de ces trois prononciations est à considérer comme orthoépique, elle sera

considérée de type plein si l'identification de cette voyelle comme /i/, /e/ ou bien /a/ ne provoque pas de doutes.

En pareilles occasions, on affaire à une variation, qui peut s'observer dans la composition phonémique de certains mots. Cette variété stimule l'établissement des certaines normes d'orthoépie.

Même un cas comme la prononciation *San Sanyč* [variante parlée du prénom et du nom plein : Alexandr Alexandrovič], dans le cas d'une réalisation complète de tous les phonèmes, doit être rattachée au type plein de diction. Cette prononciation serait de type incomplet seulement dans le cas où il serait impossible de définir par les données phonétiques complètes quel phonème a été prononcé après /n/ en *Sanyč* : /y/, /e / ou bien /a/.

En conclusion, il faut revenir à la question de l'enseignement de la prononciation d'une langue étrangère. La finalité de l'enseignement, bien évidente, doit être l'acquisition de deux types de diction dans leur interaction. Cependant, puisque il faut apprendre, tout d'abord, les rapports phonémiques, manifestés dans les unités langagières de la langue étrangère, il faut commencer par le type plein – d'autant plus que le passage vers le type incomplet peut être formulé par des règles concrètes - et non dans l'autre sens.

© Lija Bondarko, Ljudmila Verbickaja, Mirra Gordina,
Vadim Kasevič, Lev Zinder
Traduit du russe par Irina Ivanova et Jean-Baptiste Blanc

Traduit depuis l'original russe: «Stili proiznošenija i tipy proiznesenija», in: *Voprosy jazykoznanija*, N° 2, 1974, p. 64-70.

RÉFÉRENCES BIBLIOGRAPHIQUES

- AVANESOV Ruben, 1972 : *Russkoje literaturnoje proiznošenije*, Moskva. [‘La prononciation littéraire russe’]
- BONDARKO Lija, VERBICKAJA Ljudmila, 1973 : « O fonetičeskix karakteristikax zaudarnyx fleksij v sovremennom russkom jazyke », *Voprosy jazykoznanija*, 1. p. 37-49. [‘Sur les caractéristiques des flexions post-accentuées dans le russe contemporain’]
- JAKOBSON Roman, HALLE Morris, 1962 : « Fonologija i ejo otnošenije k fonetike », in : *Novoe v lingvistike*, II, Moskva., p. 231-278. [La phonologie et ses rapports à la phonétique]
- LEON Pierre R., 1971 : *Essais de phonostylistique*, *Studia Phonetica*, IV, Ottawa-Montréal : Didier.
- ŠČERBA Lev, 1937 : *Fonetika francuzskogo jazyka*, Leningrad-Moskva. [‘Phonétique du français’]
- ŠČERBA Lev, 1957 : « O raznyx stiljax proiznošenija i ob ideal’nom fonetičeskom sostave slov » in : Ščerba, *Izbrannye raboty po rusckomu jazyku*, Moskva, p. 21-25. [‘Des différentes styles de prononciation et de la composition idéale des mots’]
- ŠČERBA Lev, 1957 : « Teorija ruskogo pis’ma », in : Ščerba, *Izbrannye raboty po rusckomu jazyku*, Moskva, p. 144-179. [Théorie de l’écriture russe].
- ŠAPIRO Michael, 1968 : *Russian phonetic variants and phonostylistics*, Berkley; Los Angeles.