

# COMPLÉMENT D'UNE ÉTUDE SUR L'ACCORD DE NOMBRE NOMINAL DANS LA VARIÉTÉ URBAINE DU PORTUGAIS DE SÃO TOMÉ : STRUCTURES PRÉDICATIVES/ PASSIVES

*Silvia Figueiredo Brandão*

Universidade Federal do Rio de Janeiro/CNPq

*Paulo Vitor Lima da Gama Soares*

Universidade Federal do Rio de Janeiro/FAPERJ

L'accord de nombre (nominal et verbal) est l'un des sujets les plus discutés dans le contexte du Portugais Brésilien (ci-après PB), car non seulement c'est l'un des paramètres les plus emblématiques de la distinction entre les variétés européenne et brésilienne, mais aussi parce que, dans la dernière, il constitue une règle variable dans tous les parlers régionaux et sociaux. De plus, au Brésil, notamment en milieu urbain, le fait de ne pas faire l'accord de nombre est l'objet d'une évaluation négative de la part de certains locuteurs, qui le stigmatisent souvent car il apparaît plus fréquemment dans le discours d'individus de niveau d'éducation inférieur et/ou d'origine rurale ou rurale.

En ce qui concerne l'accord nominal en PB, depuis les premières études qui l'ont traité d'un point de vue sociolinguistique, il a été montré qu'en plus des schémas d'accord parfaits, comme (i) *os meus irmãos* (mes frères), il y a d'autres structures variables comme (ii) *aquelas coisas boa*, (ces bonnes choses), (iii) *as ideia interessante* (les idées intéressantes) et (iv) *os menino* (les garçons) / *dois menino* (deux garçons), ces derniers, des Syntagmes Nominaux (SN) de deux constituants, les plus productifs dans la modalité parlée.

Cette complexité est déterminée par des facteurs liés, d'une part, à la structure du SN et aux changements morphophonologiques résultants du mécanisme de flexion, et, d'autre part, à des facteurs liés à des caractéristiques des locuteurs (parmi lesquelles, le sexe, l'âge, le niveau d'instruction, la zone d'origine ou de résidence), à la situation d'intercommunication (formelle, informelle), et la modalité de langue (parlée, écrite).

En 2008-2009, la perspective d'observer la variation de nombre dans des variétés du portugais à profil urbain a incité les chercheurs de trois universités – celles de Lisbonne, Fédérale de Rio de Janeiro et de Madère – à mener des recherches comparatives entre les variétés brésiliennes, africaines et européennes (continentales et insulaires), en commençant par l'accord nominal et l'accord verbal à la troisième personne du pluriel. Sur la base des analyses réalisées à partir de *corpora*, pour la plupart, organisées par les membres du projet qui s'est alors constitué – Étude comparative des normes d'accord dans les variétés africaines, brésiliennes et européennes –, il a été constaté que, contrairement au Portugais Européen (ci-après PE) urbain, dans lequel la règle d'accord semble être catégorique au sein du SN (BRANDÃO, 2013) et semi-catégorique par rapport au verbe à la troisième personne du pluriel (VIEIRA ; BAZENGA, 2013), les variétés non-européennes ont des normes variables similaires, dans la mesure où la présence ou l'absence de la marque du pluriel obéit à de fortes restrictions structurelles et sociales (BRANDÃO, 2011a, 2011b, 2013, 2018 ; BRANDÃO ; VIEIRA, 2018, entre autres).

Cependant, dans le cadre de ce projet, l'accord de nombre qui s'établit (a) entre le participe passif et le sujet de la construction et (b) entre le SN ou le Syntagme Adjectival (comme prédicatifs) et (i) le sujet ou (ii) l'objet n'a pas été abordé.

Les phrases suivantes servent d'exemple pour des structures qui sont analysées dans cette étude : (a) *minhas músicas são feitas em português* (mes chansons sont faites en portugais) / *meus filho foi batizado no católico* (mes fils ont été baptisés à l'église catholique) ; (b) *elas são professoras* (ce sont des enseignantes) / *nós somos compadre e fala dialeto* (nous sommes compères et parlons un dialecte) ;

*as praias são muito bonitas* (les plages sont très belles) / *aulas naquele tempo era diferente* / les classes à cette époque étaient différentes) ; *as crianças chegaram assustadas* (les enfants sont arrivés effrayés) / *as pessoas chegaram muito cansada* (les gens sont arrivés très fatigués).

Ainsi, afin d'élargir et de compléter les connaissances sur la dynamique de l'accord, le (non)marquage du nombre pluriel dans les structures prédicatives/passives (ci-après EPPs) dans la variété urbaine du Portugais de São Tomé (ci-après PST) est analysé quantitativement et qualitativement, comparant ses résultats à ceux obtenus par Brandão (2011b) dans son analyse sur l'accord dans les constituants du SN.

São Tomé est la capitale de la République Démocratique de São Tomé et Príncipe, un archipel situé dans le golfe de Guinée, sur la côte ouest de l'Afrique et composé principalement des deux îles qui donnent son nom au pays. Avec un total de 1 001 km<sup>2</sup>, elle compte 197 700 habitants, selon le recensement de 2017 (INE). Actuellement, la majorité de la population (67,7 %) vit en zone urbaine, avec un taux d'alphabétisation d'environ 90,1 %. Dans le pays, il existe quatre établissements d'enseignement supérieur, dont un public. Dans la capitale, São Tomé, où se trouve le principal port du pays, les principales activités commerciales ont lieu, malgré sa structure précaire.

Au pays, il y a aussi, en plus de quatre créoles de base lexicale portugaise – le forro (ou santomé) et l'angolar, tous les deux sur l'île de São Tomé, le lung'ie (ou principense) sur l'île du Príncipe et le fa d'ambô (ou anobonense), sur l'île de Ano Bom (province de Guinée Équatoriale) – le portugais des Tongas, le cap-verdien (créole de base portugaise, originaire du Cap-Vert) et « vestiges de langues du groupe bantou » (HAGEMEIJER 2009, p. 1).

Ainsi, l'intérêt de l'étude du PST réside, entre autres aspects, dans l'observation, dans les normes en construction de cette variété, d'un impact possible du contact entre portugais et créoles, en particulier, avec forro, puisqu'entre autres pays africains qui ont le portugais comme langue officielle, c'est la seule variété parlée comme L1 ou L2 par la majorité de la population.

L'analyse des EPPs a été réalisée à la lumière des hypothèses théorico-méthodologiques de la Théorie de la Variation et du Changement (WEINREICH ; LABOV ; HERZOG, 1968), à partir de données recueillies d'entretiens de profil sociolinguistique réalisés en 2009 et appartenant au CORPUS VAPOR (Variétés de Portugais), du Centre de Linguistique de l'Université de Lisbonne.

Les informateurs, au nombre de 22, sont répartis par sexe/genre, trois tranches d'âge (A-18-35, B-36-55 et C-56-75 ans) et deux niveaux d'enseignement (primaire

et secondaire), deux par cellule, à l'exception du groupe plus âgé (C), dans lequel il n'y a qu'un informateur masculin avec une éducation primaire et une femme avec une éducation secondaire. La *fréquence variable d'utilisation d'un créole* (en général, forro) a également été contrôlée pour vérifier si sa fréquence plus ou moins grande impliquait une probabilité plus ou moins grande de ne pas marquer le pluriel. La variable se composait de 3 facteurs : Fréquence Ø (FØ), nulle ou faible, – les individus s'expriment principalement en portugais ; Fréquence 1 (F1), moyenne – les individus s'expriment en portugais, mais dominant un créole et en font un usage occasionnel ; Fréquence 2 (F2), élevée – les individus, bien qu'ils parlent portugais, s'expriment régulièrement en créole.

En plus des quatre variables extralinguistiques décrites ci-dessus, les neuf variables structurelles considérées par Scherre (1991) ont été contrôlées : *structure prédicative* ; *parallélisme formel des séquences de prédicats/participes dans le discours* ; *processus morphophonologiques de formation du pluriel* ; *tonicité du mot au singulier* ; *caractéristiques formelles du sujet de construction* ; *caractéristiques formelles du verbe de la construction* ; *ordre des éléments dans la structure* ; *matériel intervenant entre le verbe et le prédicatif/participe* ; *type de structure* (active, passive).

Sur les 22 entretiens, dont la durée varie de 25 à 55 minutes, seules 75 données ont été obtenues, ce qui s'explique par la faible productivité, dans les entretiens sociolinguistiques, des types de structures représentées ici. Sur les 22 informateurs, quatre n'avaient pas d'EPP dans, respectivement, 25,5 ; 23,14 ; 5,18 et 14,2 minutes d'entrevue, les plus brèves du *corpus*. En revanche, le nombre d'occurrences par individu variait d'une à dix, non pas nécessairement en fonction de la durée d'enregistrement, mais des thèmes abordés.

L'analyse des données soumises au programme Goldvarb-X a indiqué que, globalement, dans le discours populaire urbain de São Tomé, l'absence de la marque du pluriel affecte 25 des 75 occurrences d'EPPs (33,3 %), étant conditionnée par les variables *niveau d'éducation*, *caractéristiques formelles du sujet de la construction* et *fréquence d'utilisation d'un créole*.

Compte tenu, d'une part, du très petit échantillon et, d'autre part, du nombre expressif de variables, il a été nécessaire de regrouper les facteurs (comme ce fut le cas pour les deuxième et troisième variables sélectionnées), et de procéder à une analyse ponctuelle de caractère qualitatif.

Concernant le *niveau d'instruction*, également sélectionné en premier dans l'étude de Soares (2021), les individus ayant un niveau d'instruction élémentaire (17/29 données, 58,6 %, P.R. 0.695) sont ceux qui utilisent le moins le morphème

pluriel. La deuxième variable sélectionnée, de nature structurelle – *les caractéristiques formelles du sujet de la construction* – était initialement composée de 8 facteurs, soit deux de plus que ceux suggérés par Sherre (1991). L'observation attentive du nombre des données et des fréquences de marquage et de non-marquage du pluriel au sein de ce groupe a permis de réduire les facteurs à seulement deux : sujet explicite marqué et sujet (non) explicite non marqué. On a vérifié que le premier d'entre eux était le plus actif pour l'annulation de la marque du pluriel (20/39 données, 51,3%, PR .751), ce qui a prouvé l'hypothèse initiale selon laquelle plus le sujet (élément à sa gauche) présentait une marque plurielle morphologique ou sémantique plus grande serait la tendance à annuler la marque numérique dans les EPPs.

Le croisement des deux variables évoquées est également instructif quant au fait que les locuteurs les moins instruits sont ceux qui conduisent l'annulation du morphème du pluriel s'il est déjà explicite à gauche (13/16 occurrences, 81%), ce qui est dans la lignée de ce qui est observé dans le SN, dans lequel la marque du pluriel a son contexte par excellence dans le *locus* gauche, étant moins fréquente plus à droite (ex : *os menino atrevido* (les garçons insolents).

Une autre variable qui s'est révélée significative dans les analyses du PST est la *fréquence d'utilisation d'un créole*, composée de trois facteurs qui cherchent à évaluer le contact de l'individu avec d'autres langues, notamment le forro. Tous les informateurs qui ont présenté des données de EPPs et ayant déclaré qu'ils parlent essentiellement le portugais (F0), ont appliqué la marque du pluriel (24 occurrences). La variation se retrouve chez les individus qui, bien que lusophones, utilisent occasionnellement un créole (F1) ou s'expriment régulièrement dans l'une de ces langues (F3).

Pour compléter l'analyse des EPPs, basée sur un petit nombre de données (75), une nouvelle observation de l'étude de Brandão (2011b) a été réalisée par l'informateur, prenant cette fois en compte chaque SN dans son ensemble par nombre de constituants et en indiquant ceux avec au moins une marque de non-application du morphème du pluriel. Il a été constaté que le *corpus* de cette étude se composait d'un total de 1.251 SNs, dont 1.061 (84%) de deux constituants (SN2), 168 (13,4%) de trois (SN3) et 22 de quatre ou cinq (SN4/5) : 1,75%.

Dans l'article, on présente, en deux tableaux, les index d'absence du morphème du pluriel en EPPs et SNs (selon le nombre de leurs constituants) dans le discours des 11 informateurs de niveau élémentaire (tableau 7) et 11 de niveau secondaire.

L'observation de chaque tableau démontre, dans leurs totaux partiels, que, dans l'une ou l'autre des structures abordées, les individus ayant un niveau d'éducation

fondamental sont ceux qui se disputent le plus la non-application du pluriel, soit dans le cadre des EPPs (58,6%), ou dans les SNs (SN-2 : 41,7% ; SN3 : 84,6% ; SN 4/5 = 87,5%).

Parmi les informateurs de niveau intermédiaire, dont les EPPs présentent 17,4% d'annulation de la marque du pluriel, les taux liés aux SNs conservent la même échelle d'annulation progressive en fonction du nombre croissant de constituants du SN : SN2 : 5,9% ; SN3 : 11,5% ; SN4/5 : 26%.

En revanche, en ce qui concerne spécifiquement les EPPs : (a) un seul individu de niveau fondamental a fait le pluriel dans les 3 occurrences avec lesquelles il a contribué à l'échantillon ; (b) parmi les 10 informateurs de niveaux intermédiaires qui ont présenté d'occurrences de EPPs, huit ont appliqué la marque du pluriel de façon catégorique (dont un, y compris, dans tous les types de SN) et seulement deux l'ont annulée - ST-A-2-ha (50% ) et ST-C-2-hb (42,9 %), avec des indices très similaires.

En conclusion, il a été possible de vérifier que, dans la variété de São Tomé, le marquage du pluriel est variable, corroborant les résultats obtenus dans l'analyse au sein du SN, dans laquelle des facteurs structurels et sociaux sont actifs. Parmi ces derniers, se distinguent ceux liés au niveau d'instruction et à la fréquence d'utilisation d'un créole (première et troisième variables retenues dans l'analyse variationniste). Il est clair que les individus qui ont le moins fréquenté l'école sont ceux qui appliquent le moins la marque du pluriel, de même que ceux qui ont le plus de contacts avec un créole.

Quant aux conditionnements structurels, *les caractéristiques formelles du sujet de la construction*, retenues en deuxième position dans l'étude de Scherre (1991), sans toutefois observer l'action du parallélisme formel, se sont avérées pertinentes. La réorganisation des facteurs de la variable indique qu'il y a une plus grande probabilité de ne pas appliquer le morphème du pluriel lorsque le sujet est explicite et présente une marque sémantique et/ou morphologique de pluralité.

Les analyses sur l'accord dans le SN dans les variétés africaines du portugais ont montré un comportement similaire à celui observé dans différentes variétés du PB, en ce qui concerne les normes observées et les variables qui les conditionnent. Il est à noter que ceux de nature sociale (notamment *niveau d'éducation*) sont toujours très significatives, décrivant ce que, en relation au PB, Lucchesi (2015, entre autres travaux) a appelé la polarisation sociolinguistique, reflet du contact multilingue qui s'est produit (et se produit encore) dans la phase de colonisation.

Bien que le standard de référence du PST soit, en termes officiels, le PE, dans lequel la règle d'accord dans le SN, dans la variété urbaine, est catégorique,

comme le suggère Brandão (2013), le standard santoméen naissant s'en écarte, certainement en fonction de l'histoire sociolinguistique du pays, marquée par le contact multilingue, dans laquelle le forro joue certainement un rôle important.

Il y a des indications (HAGEMEIJER, 2009, 2018 ; ARAUJO, 2020), sur la façon dont les normes qui régissent l'utilisation du portugais à São Tomé sont en train d'être définies (apparemment, de manière accélérée) au détriment des créoles, en particulier du forro, la langue autrefois plus répandue et caractérisée par certains informateurs de cette recherche comme « la langue nationale ». Hagemeyer (2009), qui avait déjà commenté la diversité des normes qui coexistent dans le PST, caractérise la situation actuelle (2018, p. 182) comme assez différente de celle d'un « passé récent dans lequel le portugais était L2 pour la plupart des habitants des îles. »

Malgré le peu de (mauvaises) données concernant les EPPs, on a cherché à en faire le meilleur usage possible comme contribution à la connaissance de la dynamique du (non)marquage du pluriel dans le PST et, par extension, dans les autres variétés non européennes du portugais qui, en ce qui concerne les thèmes abordés ici, semblent présenter des convergences marquantes, parmi lesquelles celles concernant le contact multilingue, l'un des facteurs, sans doute, déterminants dans la construction de ses normes.

## RÉFÉRENCES

ARAUJO, Gabriel Antunes. Portuguese language expansion in São Tomé and Príncipe: an overview. *Diadorim*, Rio de Janeiro, v. 22, n° 1, p. 57-78, 2020.

BRANDÃO, Silvia Figueiredo. Concordância nominal em duas variedades do português: convergências e divergências. *Revista Veredas*, v. 15, n. 1, p. 164-178, 2011a.

BRANDÃO, Silvia Figueiredo. *O cancelamento da marca de número nominal na variedade urbana não standard do Português de São Tomé*. Comunicação apresentada ao XVI Congreso Internacional de la Asociación de Lingüística y Filología de la América Latina (ALFAL). Alcalá de Henares, 6-9 de junho, 2011b.

BRANDÃO, Silvia Figueiredo. Patterns of agreement within the Noun Phrase. *Journal of Portuguese Linguistics*, v. 12, n. 2, p. 51-100, 2013.

BRANDÃO, Silvia Figueiredo. Concordância nominal no Português de São Tomé e no Português de Moçambique. In: BRANDÃO, Silvia Figueiredo.

*Duas variedades africanas do Português: variáveis fonético-fonológicas e morfossintáticas.* São Paulo: Blucher, 2018. p. 203-244.

BRANDÃO, Silvia Figueiredo; VIEIRA, Silvia Rodrigues. The agreement continuum in urban samples of African, Brazilian and European varieties of Portuguese. In: LÓPEZ, Laura Álvarez; GONÇALVES, Perpétua; AVELAR, Juanito Ornelas de. *The Portuguese language continuum in Africa and Brazil.* Amsterdam/Philadelphia: John Benjamins Publishing Company, 2018. p. 267-289.

HAGEMEIJER, Tjerk. As línguas de São Tomé e Príncipe. *Revista de Crioulos de Base Lexical Portuguesa e Espanhola*, Macau, v. 1, n. 1, p. 1-27, 2009.

HAGEMEIJER, Tjerk. From creoles to Portuguese: Language shift in São Tomé and Príncipe. In: LÓPEZ, Laura Álvarez; GONÇALVES, Perpétua; AVELAR, Juanito Ornelas de. *The Portuguese language continuum in Africa and Brazil.* Amsterdam/Philadelphia: John Benjamins Publishing Company, 2018. p. 169-184.

INSTITUTO NACIONAL DE ESTATÍSTICA –INE, São Tomé e Príncipe. Disponível em: <https://www.ine.st/>. Acesso em: 15 abr. 2021.

LUCCHESI, Dante. *Língua e sociedade partidas: a polarização sociolinguística do Brasil.* São Paulo: Contexto, 2015.

SCHERRE, Maria Marta Pereira. A concordância de número nos predicativos e nos participípios passivos. *Organon*, Porto Alegre, v. 5, n. 18, p. 52-70, 1991.

SOARES, Paulo Vitor Lima da Gama. *Concordância de número em estruturas predicativas/passivas no Português de São Tomé: resultados preliminares.* Comunicação apresentada à 42ª Jornada Giulio Massarani de Iniciação Científica, Tecnológica, Artística e Cultural. Rio de Janeiro, Universidade Federal do Rio de Janeiro, março de 2021.

VIEIRA, Silvia Rodrigues; BAZENGA, Aline. Patterns of third person verbal agreement. *Journal of Portuguese Linguistics*, Lisbon, v. 12, n. 2, p. 7-50, 2013.

WEINREICH, Uriel; LABOV, William; HERZOG, Martin. Empirical foundations for a theory of linguistic change. In: LEHMANN, Winfred.; MALKIEL, Yakov. *Directions for historical linguistics.* Austin: University of Texas Press, 1968. p. 97-195.